

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-005949

(43)Date of publication of application : 10.01.2003

(51)Int.Cl.

G06F 3/16
G10L 11/00
G10L 11/02
G10L 15/00
G10L 15/02
G10L 15/04

(21)Application number : 2001-186926

(71)Applicant : NEC CORP

(22)Date of filing : 20.06.2001

(72)Inventor : YAMADA EIKO
HAGANE HIROSHI
YOSHIDA KAZUNAGA

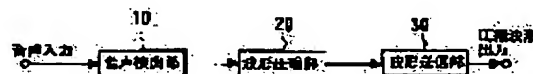
(54) SERVER CLIENT TYPE VOICE RECOGNIZING DEVICE AND METHOD

(57)Abstract:

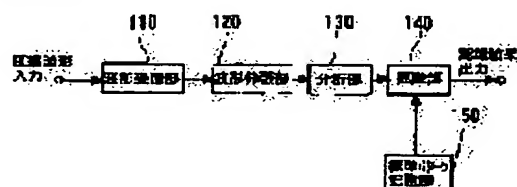
PROBLEM TO BE SOLVED: To provide a voice recognizing device capable of shortening a transmission time, and reducing costs.

SOLUTION: A terminal side equipment is provided with a voice detecting part 10 for detecting the voice block of inputted data, a waveform compressing part 20 for compressing the waveform data of the voice block detected by the voice detecting part 10, and a waveform transmitting part 30 for outputting the waveform data compressed by the waveform compressing part 20. A server side device is provided with a waveform receiving part 110 for receiving waveform data outputted from the terminal side equipment, a waveform extending part 120 for extending the waveform data received by the waveform receiving part 110, an analyzing part 130 for analyzing the waveform data extended by the waveform extending part 120, and a recognizing part 140 for performing recognition processing, and for outputting the recognized result.

(a)
【請求項】



(b)
【発明】



LEGAL STATUS

[Date of request for examination]

12.02.2003

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the
examiner's decision of rejection or application converted
registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of
rejection]

[Date of requesting appeal against examiner's decision of
rejection]

[Date of extinction of right]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2003-5949

(P2003-5949A)

(43)公開日 平成15年1月10日(2003.1.10)

| (51)Int.Cl. | 識別記号 | F I | テマコード(参考) |
|---------------|-------|--------------|-------------------|
| G 0 6 F 3/16 | 3 4 0 | G 0 6 F 3/16 | 3 4 0 A 5 D 0 1 5 |
| G 1 0 L 11/00 | | G 1 0 L 3/00 | 5 5 1 A |
| 11/02 | | | 5 1 3 C |
| 15/00 | | | 5 1 5 A |
| 15/02 | | | |

審査請求 未請求 請求項の数27 O L (全 24 頁) 最終頁に続く

(21)出願番号 特願2001-186926(P2001-186926)

(22)出願日 平成13年6月20日(2001.6.20)

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 山田 栄子

東京都港区芝五丁目7番1号 日本電気株式会社内

(72)発明者 羽金 広

東京都港区芝五丁目7番1号 日本電気株式会社内

(74)代理人 100080816

弁理士 加藤 朝道

最終頁に続く

(54)【発明の名称】 サーバ・クライアント型音声認識装置及び方法

(57)【要約】

【課題】送信時間の短縮、コストの低減を図る音声認識装置の提供。

【解決手段】端末側の装置が、入力されたデータの音声区間を検出する音声検出部10と、音声検出部10で検出された音声区間の波形データを圧縮する波形圧縮部20と、波形圧縮部20で圧縮された波形データを出力する波形送信部30と、を備え、前記サーバ側の装置が、端末側の装置から出力された波形データを受信する波形受信部110と、波形受信部110で受信された波形データを伸張する波形伸張部120と、波形伸張部120で伸張された波形データを分析する分析部130と、認識処理を行い、認識した結果を出力する認識部140と、を備えている。

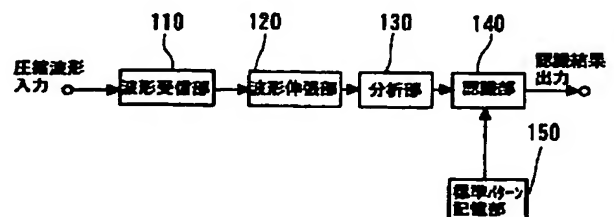
(a)

【端末側】



(b)

【サーバ側】



【特許請求の範囲】

【請求項 1】 端末側の装置と、サーバ側の装置と、を備えた音声認識装置であって、
前記端末側の装置が、
入力されたデータの音声区間を検出する音声検出部と、
前記音声検出部で検出された音声区間の波形データを圧縮する波形圧縮部と、
前記波形圧縮部で圧縮された波形データを出力する波形送信部と、
を備え、
前記サーバ側の装置が、
前記端末側の装置から出力された波形データを受信する波形受信部と、
前記波形受信部で受信された波形データを伸張する波形伸張部と、
前記波形伸張部で伸張された波形データを用いて認識処理を行い、認識した結果を出力する認識部と、
を備えている、ことを特徴とする音声認識装置。

【請求項 2】 端末側の装置と、サーバ側の装置と、を備えた音声認識装置であって、
前記端末側の装置が、
入力された音声の波形データと、前記サーバ側の装置から送信される波形データ再送要求信号とを受信する波形・信号受信部と、
前記波形・信号受信部で受信された波形データの音声区間を検出する音声検出部と、
前記音声検出部で検出された音声区間の波形データを圧縮する波形圧縮部と、
波形データを出力する波形送信部と、
前記波形圧縮部で圧縮された波形データを一時記憶しながら前記波形データを、前記波形送信部へ送信すると共に、前記波形・信号受信部において前記サーバ側の装置からの波形データ再送要求信号が受信された場合に、記憶された波形データを、前記波形送信部へ送信する波形記憶部と、
を備え、
前記サーバ側の装置が、
前記端末側の装置から出力された波形データを受信する波形受信部と、
前記波形受信部で受信された波形データを伸張する波形伸張部と、
前記波形伸張部で伸張された波形データを用いて認識処理を行い、認識した結果を出力する認識部と、
前記波形受信部で圧縮波形データの受信に失敗した場合には、波形データの再送要求信号を前記波形受信部から受け取り波形データ再送要求信号を出力する波形データ再送要求信号送信部と、を備えている、ことを特徴とする音声認識装置。

【請求項 3】 端末側の装置と、サーバ側の装置と、を備えた音声認識装置であって、

前記端末側の装置が、
入力された音声の波形データと、前記サーバ側の装置から送信される波形データ再送要求信号とを受信する波形・信号受信部と、
前記波形・信号受信部で受信された波形データの音声区間を検出する音声検出部と、
前記音声検出部で検出された音声区間の波形データを圧縮する波形圧縮部と、波形データを出力する波形送信部と、

10 前記波形圧縮部で圧縮された波形データを一時記憶しながら前記波形データを、前記波形送信部へ送信すると共に、前記波形・信号受信部において、前記サーバ側の装置からの波形データ再送要求信号が受信された場合に、記憶された波形データを、前記波形送信部へ送信する波形記憶部と、
前記音声検出部で音声の検出後に該検出がキャンセルされた場合、前記音声検出部から送信された始端キャンセル時の信号を受け取り始端キャンセル信号を出力する始端キャンセル信号送信部と、

20 を備え、
前記サーバ側の装置が、
前記端末側の装置から出力された波形データと、前記始端キャンセル信号とを受信する波形・信号受信部と、
前記波形・信号受信部で受信された音声の波形データを伸張する波形伸張部と、
前記波形伸張部で伸張された波形データを用いて認識処理を行い、認識した結果を出力する認識部と、
前記波形受信部で圧縮波形データの受信に失敗した場合には、波形データの再送要求信号を前記波形受信部から受け取り波形データ再送要求信号を出力する波形データ再送要求信号送信部と、
30 を備え、前記認識部は、前記波形・信号受信部において始端キャンセル信号を受信した場合に、認識処理を中止する、ことを特徴とする音声認識装置。

【請求項 4】 端末側の装置と、サーバ側の装置と、を備えた音声認識装置であって、
前記端末側の装置が、
入力された波形データと、前記サーバ側の装置から送信される波形データ再送要求信号と、前記サーバ側の装置から送信される前記サーバ側の装置で利用可能な圧縮方式情報とを少なくとも受信する波形・信号・圧縮方式受信部と、
前記波形・信号・圧縮方式受信部で受信された波形データの音声区間を検出する音声検出部と、
前記音声検出部で検出された音声区間の波形データを圧縮する波形圧縮部と、波形データを出力する波形送信部と、
前記波形圧縮部で圧縮された波形データを一時記憶しながら前記波形データを、前記波形送信部へ送信すると共に、前記波形・信号・圧縮方式受信部において、前記サ

一バ側の装置からの波形データ再送要求信号が受信された場合に、記憶された波形データを、前記波形送信部へ送信する波形記憶部と、

前記波形・信号・圧縮方式受信部において前記サーバ側で利用可能な圧縮方式情報を受信した場合に、前記波形・信号・圧縮方式受信部から送られた圧縮方式情報より最適な圧縮方式を選択する圧縮方式選択部と、前記圧縮方式選択部で選択された圧縮方式のインデックスを作成する圧縮方式インデックス作成部と、を備え、

前記波形圧縮部は、前記音声検出部で検出された音声区間の波形データを圧縮し、前記圧縮方式インデックス作成部で作成された圧縮方式インデックスを波形データの一部に組み込み、

前記サーバ側の装置が、

前記端末側の装置から送信された波形データと、圧縮方式要求信号とを受信する波形・信号受信部と、

前記波形・信号受信部で受信された音声の波形データを伸張する波形伸張部と、

前記波形伸張部で伸張された波形データを用いて認識処理を行い、認識した結果を出力する認識部と、

前記波形・信号受信部で圧縮波形データの受信に失敗した場合には、波形データの再送要求信号を前記波形受信部から受け取り波形データ再送要求信号を出力する波形データ再送要求信号送信部と、

前記サーバ側の装置で利用可能な圧縮方式の情報を記憶する圧縮方式記憶部と、

前記波形・信号・圧縮方式受信部から圧縮方式要求信号が送られた場合に前記圧縮方式記憶部に記憶されている圧縮方式情報を取得し、前記端末側の装置に圧縮方式情報を出力する圧縮方式取得部と、

前記波形伸張部で伸張されたデータから圧縮方式のインデックスを取得する圧縮方式インデックス取得部と、

前記圧縮方式インデックス取得部で取得された圧縮方式のインデックスから認識エンジンを選択する認識エンジン選択部と、

前記認識エンジン選択部で選択されたエンジンを記憶されているエンジンの中から設定する認識エンジン設定部と、

を備えている、ことを特徴とする音声認識装置。

【請求項 5】前記端末側の装置が、

合成音を合成する音声合成部と、

前記音声合成部で合成された合成音の情報を作成し、合成音を出力する合成音情報作成部と、

を備え、

前記波形圧縮部は、前記音声検出部で検出された音声区間の波形データを圧縮し、前記合成音情報作成部で作成された合成音情報を前記波形データの一部に組み込み、前記サーバ側の装置が、

前記波形伸張部で伸張されたデータから合成音情報を取

得する合成音情報取得部を備え、前記認識部が、認識が終了した場合に前記合成音情報取得部で取得された合成音情報より合成音と認識結果との対応付けを行い、対応付けされた認識結果、又は認識結果と合成音情報を出力する、ことを特徴とする請求項 3 又は 4 に記載の音声認識装置。

【請求項 6】前記サーバ側の装置において、前記波形圧縮部は、前記音声検出部で検出された音声区間の波形データを圧縮し、前記合成音情報作成部で作成された合成音情報と、前記圧縮方式インデックス作成部で作成された圧縮方式インデックスを波形データの一部に組み込む、ことを特徴とする請求項 5 に記載の音声認識装置。

【請求項 7】前記サーバ側装置において、前記波形・信号受信部に代わって、前記端末側の装置から送信された波形データと、始端キャンセル信号と、圧縮方式要求信号と、コンテンツ側から送信されたタスク情報と、を受信する波形・信号・タスク情報受信部を備え、

前記波形・信号・タスク情報受信部で受信されたタスク情報を記憶するタスク情報記憶部と、

タスク情報と該タスク使用時に利用可能な 1 個以上の圧縮方式を記憶する圧縮方式・タスク情報対応テーブル記憶部と、

前記波形・信号・タスク情報受信部において圧縮方式要求信号を受信した場合に、前記タスク情報記憶部から送られたタスク情報と前記圧縮方式・タスク情報対応テーブル記憶部から送られたタスク情報と圧縮方式との対応テーブルより利用可能な圧縮方式情報を取得し、前記端末側の装置に出力する圧縮方式取得部とを備えている、ことを特徴とする請求項 4 又は 5 に記載の音声認識装置。

【請求項 8】前記端末側の装置において、

前記波形・信号・圧縮方式受信部に代わって、入力された波形データと、コンテンツ側から送信されたタスク情報と、前記サーバ側の装置から送信された波形データ再送要求信号と、前記サーバ側の装置から送信された前記サーバ側の装置で利用可能な圧縮方式情報とを受信する波形・信号・圧縮方式・タスク情報受信部を備え、

前記波形・信号・圧縮方式・タスク情報受信部で受信されたタスク情報を記憶するタスク情報記憶部と、

タスク情報とそのタスク使用時に利用可能な 1 個以上の圧縮方式の対応テーブルを記憶する圧縮方式・タスク情報対応テーブル記憶部と、

前記波形・信号・圧縮方式・タスク情報受信部においてサーバ側で利用可能な圧縮方式情報を受信した場合に、前記タスク情報記憶部から送られたタスク情報と、前記圧縮方式・タスク情報対応テーブル記憶部から送られたタスク情報と圧縮方式との対応テーブルと、前記波形・信号・圧縮方式・タスク情報受信部から送られた前記サーバ側の装置で利用可能な圧縮方式情報とをともに、最適な

圧縮方式を選択する圧縮方式選択部と、
を備えたことを特徴とする請求項 5 に記載の音声認識装置。

【請求項 9】 端末から出力された圧縮波形データを受信して伸張し、該伸張された波形データを用いて認識処理を行い、認識した結果を出力するサーバ装置に接続され、サーバ・クライアント音声認識装置を構成する端末であって、

入力されたデータの音声区間を検出する音声検出部と、
前記音声検出部で検出された音声区間の波形データを圧縮する波形圧縮部と、

前記波形圧縮部で圧縮された波形データを出力する波形送信部と、

を備えている、ことを特徴とする端末。

【請求項 10】 入力されたデータの音声区間を検出し、該検出された音声区間の波形データを圧縮する波形圧縮部と、前記波形圧縮部で圧縮された波形データを出力する端末に接続され、前記端末とともに、サーバ・クライアント音声認識装置を構成するサーバ装置であって、前記端末から出力された波形データを少なくとも受信する受信部と、

前記波形受信部で受信された波形データを伸張する波形伸張部と、

前記波形伸張部で伸張された波形データを用いて認識処理を行い、認識した結果を出力する認識部と、
を備えている、ことを特徴とするサーバ装置。

【請求項 11】 入力された音声の波形データと、前記サーバ装置から送信された波形データ再送要求信号を受信する波形・信号受信部と、

前記波形・信号受信部で受信された波形データの音声区間を検出する音声検出部と、

前記音声検出部で検出された音声区間の波形データを圧縮する波形圧縮部と、

波形データをサーバ装置に送信する波形送信部と、

前記波形圧縮部で圧縮された波形データを、一時記憶しながら、前記波形データを、前記波形送信部へ送信すると共に、前記波形・信号受信部において前記サーバ装置からの波形データ再送要求信号が受信された場合に、記憶された波形データを、前記波形送信部へ送信する波形記憶部と、

をさらに備えている、ことを特徴とする請求項 9 に記載の端末。

【請求項 12】 前記受信部で、圧縮波形データの受信に失敗した場合に、波形データ再送要求信号を前記端末に送信する波形データ再送要求信号送信部を備えている、ことを特徴とする請求項 10 に記載のサーバ装置。

【請求項 13】 前記音声検出部で音声の検出後に該検出がキャンセルされた場合、前記音声検出部から送信された始端キャンセル時の信号を受け取り、始端キャンセル信号を前記サーバ装置に送信する始端キャンセル信号送

信部をさらに備えている、ことを特徴とする請求項 11 に記載の端末。

【請求項 14】 前記端末において音声の検出後に該検出がキャンセルされた場合に出力される始端キャンセル信号を、前記受信部が受信した場合に、前記受信部からの通知に基づき、前記認識部は、認識処理を中止する、ことを特徴とする請求項 10 に記載のサーバ装置。

【請求項 15】 前記端末の波形・信号受信部が、前記サーバ側の装置から送信された前記サーバ側の装置で利用可能な圧縮方式情報を受信した場合に、前記圧縮方式情報より最適な圧縮方式を選択する圧縮方式選択部と、前記圧縮方式選択部で選択された圧縮方式のインデックスを作成する圧縮方式インデックス作成部と、を備え、前記波形圧縮部は、前記音声検出部で検出された音声区間の波形データを圧縮し、前記圧縮方式インデックス作成部で作成された圧縮方式インデックスを波形データの一部に組み込む、ことを特徴とする請求項 11 又は 13 に記載の端末。

【請求項 16】 前記受信部は、前記端末から送信される圧縮方式要求信号を受信し、

前記サーバ側で利用可能な圧縮方式の情報を記憶する圧縮方式記憶部と、

前記受信部で圧縮方式要求信号を受信した場合に、前記圧縮方式記憶部に記憶されている圧縮方式情報を取得し、前記端末側に圧縮方式情報を出力する圧縮方式取得部と、

前記波形伸張部で伸張されたデータから圧縮方式のインデックスを取得する圧縮方式インデックス取得部と、

前記圧縮方式インデックス取得部で取得された圧縮方式のインデックスから認識エンジンを選択する認識エンジン選択部と、

前記認識エンジン選択部で選択されたエンジンを記憶されているエンジンの中から設定する認識エンジン設定部と、

を備えている、ことを特徴とする請求項 10、12、14 のいずれかに記載のサーバ装置。

【請求項 17】 合成音を合成する音声合成部と、前記音声合成部で合成された合成音の情報を作成し、合成音を出力する合成音情報作成部と、

を備え、前記波形圧縮部は、前記音声検出部で検出された音声区間の波形データを圧縮し、前記合成音情報作成部で作成された合成音情報を前記波形データの一部に組み込む、ことを特徴とする請求項 11、13、15 のいずれかに記載の端末。

【請求項 18】 前記波形伸張部で伸張されたデータから合成音情報を取得する合成音情報取得部を備え、前記認識部が、認識が終了した場合に前記合成音情報取得部で取得された合成音情報より合成音と認識結果との対応付けを行い、対応付けされた認識結果、又は認識結果と合成音情報を出力する、ことを特徴とする請求項 10、1

2、14、16のいずれかに記載のサーバ装置。

【請求項19】前記受信部が、前記端末側の装置から送信される波形データと、始端キャンセル信号と、圧縮方式要求信号と、コンテンツ側から送信されたタスク情報と、を受信し、

前記波形・信号・タスク情報受信部で受信されたタスク情報を記憶するタスク情報記憶部と、

タスク情報と該タスク使用時に利用可能な1個以上の圧縮方式を記憶する圧縮方式・タスク情報対応テーブル記憶部と、

前記波形・信号・タスク情報受信部において圧縮方式要求信号を受信した場合に、前記タスク情報記憶部から送られたタスク情報と前記圧縮方式・タスク情報対応テーブル記憶部から送られたタスク情報と圧縮方式との対応テーブルより利用可能な圧縮方式情報を取得し、前記端末側の装置に出力する圧縮方式取得部とを備えている、ことを特徴とする請求項10、12、14、16、18のいずれかに記載のサーバ装置。

【請求項20】前記波形・信号受信部では、入力された波形データと、コンテンツ側から送信されたタスク情報と、前記サーバ側の装置から送信される波形データ再送要求信号と、前記サーバ側の装置から送信された前記サーバ側の装置で利用可能な圧縮方式情報とを受信し、前記受信されたタスク情報を記憶するタスク情報記憶部と、

前記タスク情報とそのタスク使用時に利用可能な1個以上の圧縮方式の対応テーブルを記憶する圧縮方式・タスク情報対応テーブル記憶部と、

前記波形・信号受信部で前記サーバ側で利用可能な圧縮方式情報を受信した場合に、前記タスク情報記憶部から送られたタスク情報と、前記圧縮方式・タスク情報対応テーブル記憶部から送られたタスク情報と圧縮方式との対応テーブルと、前記波形・信号・圧縮方式・タスク情報受信部から送られた前記サーバ側の装置で利用可能な圧縮方式情報をもとに最適な圧縮方式を選択する圧縮方式選択部と、

を備えている、ことを特徴とする請求項11、13、15、17のいずれかに記載の端末。

【請求項21】サーバ装置と端末を有するサーバ・クライアントシステムの音声認識方法であって、

前記端末において、

入力されたデータの音声区間を検出するステップと、

前記検出された音声区間の波形データを圧縮するステップと、

前記圧縮された波形データをサーバに送信するステップと、

を有し、

前記サーバ装置において、

前記端末から出力された波形データを受信するステップと、

前記受信された波形データを伸張するステップと、

前記伸張された波形データを用いて認識処理を行い、認識した結果を出力するステップと、

を有する、ことを特徴とする音声認識方法。

【請求項22】サーバ装置と端末を有するサーバ・クライアントシステムの音声認識方法であって、

前記端末において、

入力された音声の波形データを受信するステップと、

前記受信された波形データの音声区間を検出するステップと、

10

前記検出された音声区間の波形データを圧縮するステップと、

前記圧縮された波形データを波形記憶部に一時記憶しながら前記サーバ装置に送信するステップと、

前記サーバ装置から送信される波形データ再送要求信号を受信した場合に、波形記憶部に記憶された波形データを、前記サーバ装置に送信するステップと、

を有し、

前記サーバ装置において、

20

前記端末から出力される波形データを受信するステップと、

前記受信された波形データを伸張するステップと、

前記伸張された波形データを用いて認識処理を行い、認識した結果を出力するステップと、

前記端末から送信される圧縮波形データの受信に失敗した場合には、波形データ再送要求信号を前記端末に出力するステップと、

を有する、ことを特徴とする音声認識方法。

【請求項23】前記端末において、

30

音声の検出後に該検出がキャンセルされた場合、始端キャンセル信号を前記サーバ装置に送信するステップをさらに有し、

前記サーバ装置において、

前記端末からの前記始端キャンセル信号を受信した場合に、認識処理を中止するステップを有する、ことを特徴とする請求項21又は22記載の音声認識方法。

【請求項24】前記端末において、

前記サーバ装置から送信された前記サーバ側で利用可能な圧縮方式情報を受信するステップと、

40

前記受信した圧縮方式情報により最適な圧縮方式を選択するステップと、

前記選択された圧縮方式のインデックスを作成するステップと、

音声区間の波形データを圧縮し、作成された圧縮方式インデックスを波形データの一部に組み込んで、前記サーバ装置に送信するステップと、

を有し、

前記サーバ装置において、

前記端末から送信される圧縮方式要求信号を受信した場合に、前記サーバ側で利用可能な圧縮方式の情報を記憶

50

する圧縮方式記憶部に記憶されている圧縮方式情報を取得し、前記端末に圧縮方式情報を出力するステップと、伸張されたデータから圧縮方式のインデックスを取得するステップと、

前記取得された圧縮方式のインデックスから認識エンジンを選択するステップと、

前記選択されたエンジンを記憶されているエンジンの中から設定するステップと、

を有する、ことを特徴とする請求項 2 2 又は 2 3 記載の音声認識方法。

【請求項 2 5】前記端末において、

合成音を合成するステップと、

前記合成された合成音の情報を作成し、合成音を出力するステップと、

検出された音声区間の波形データを圧縮し、前記合成音情報作成部で作成された合成音情報を前記波形データの一部に組み込んで、前記サーバ装置に送信するステップと、

を含み、

前記サーバ装置において、

前記伸張されたデータから合成音情報を取得するステップを有し、

認識が終了した場合に、前記取得された合成音情報より合成音と認識結果との対応付けを行い、対応付けされた認識結果、又は認識結果と合成音情報を出力する、ことを特徴とする請求項 2 2 乃至 2 4 のいずれかに記載の音声認識方法。

【請求項 2 6】前記サーバ装置において、

コンテンツ側から送信されたタスク情報を受信してタスク情報記憶部に記憶するステップと、

圧縮方式要求信号を受信した場合に、タスク情報と、圧縮方式との対応テーブルより、利用可能な圧縮方式情報を取得し、前記端末に送信するステップを有することを特徴とする請求項 2 2 乃至 2 5 のいずれかに記載の音声認識方法。

【請求項 2 7】前記端末において、

コンテンツ側から送信されたタスク情報と、前記サーバ装置から送信された前記サーバ側で利用可能な圧縮方式情報とを受信するステップと、

前記サーバ側で利用可能な圧縮方式情報を受信した場合に、前記タスク情報と、圧縮方式・タスク対応テーブル記憶部から送られたタスク情報と圧縮方式との対応テーブルと、サーバ側の装置で利用可能な圧縮方式情報をもとに最適な圧縮方式を選択するステップと、

を有する、ことを特徴とする請求項 2 6 記載の音声認識方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声認識技術に関し、特に、携帯電話等の端末（クライアント）側で音声

検出を行い、サーバ側で音声分析、認識を行うサーバ・クライアント型の音声認識装置と方法に関する。

【0002】

【従来の技術】従来より、端末側（クライアント側）で音声検出を行い、検出後の波形データをサーバ側へ送信し、サーバ側にて分析、認識処理を行うサーバ・クライアント型音声認識装置として、Dialogic CSP(Continuous Speech Processing)を用いた音声認識装置が知られている。

10 【0003】また、端末側で音声検出を行い、検出後の波形データを圧縮し、圧縮した波形データをサーバ側へ送信する方式としてVoIP(Voiceover Internet Protocol)が知られており、文献 1 (Nikkei Internet Technology; pp. 75~93, May 1998) に詳しく述べられている。またVoIPで送信された波形データを伸張し、再度認識用の音声検出を行い、検出後の波形データを分析、認識処理を行うサーバ・クライアント型音声認識装置も知られている。

20 【0004】また、近年、3GPP(Third Generation Partnership Project)のワークグループの1つであるETSI(the European Telecommunications Standards Institute)-STQ Aurora DSR(Distributed Speech Recognition) Working Groupが進めている標準化プロジェクトにおいて、端末側で音声検出、分析を行い、分析後のパラメータ（特徴ベクトル）をサーバ側に送信し、サーバ側で音声認識を行うサーバ・クライアント型音声認識装置が提案されている。

30 【0005】従来のサーバ・クライアント型音声認識装置に用いられているDialogic CSPでは、端末側で検出された波形データを圧縮せずそのままサーバへ送信しているため、送信時の時間、コストを多く必要とする、という問題点を有している。

40 【0006】また、VoIPを応用した方式では、端末側で伸張し音声検出を行ったデータに対し、サーバ側で認識用の音声検出を再度行っているため、重複した音声検出処理が無駄な処理となっている。更に、音声認識用の音声検出をサーバ側で行っていることから、サーバ側で短い音声について音声の始端検出がキャンセルされた場合に、端末側で音声検出キャンセル情報を受信するのが遅れ、それによってアプリの動作に遅延が生じるという、という問題点を有している。

50 【0007】また、ETSI-STQ Aurora DSR Working Groupが進めているサーバ・クライアント型音声認識装置の枠組みについては、分析後のパラメータが予め決められているため、独自仕様のパラメータを使用できないという欠点や、分析部が端末側に設定されているため、新規の分析手法を端末側に搭載するためのコスト、時間を多く必要とするという、という問題点を有している。

【0008】

【発明が解決しようとする課題】端末側（クライアント

側)で音声検出を行い、検出後の波形データをサーバ側へ送信し、サーバ側にて分析、認識処理を行う構成の従来のサーバ・クライアント型音声認識装置は、端末側で検出された波形データを圧縮することなく、そのままサーバへ送信しているため、送信時の時間、コストを多く必要とするという、問題点を有している。

【0009】また、端末側で音声検出を行い、検出後の波形データを圧縮し、圧縮した波形データをサーバ側へ送信し、サーバ側にて波形データを伸張し、再度、認識用の音声検出を行い、検出後の波形データを分析、認識処理を行う構成の従来のサーバ・クライアント型音声認識装置は、端末側で伸張し音声検出を行ったデータに対し、サーバ側で認識用の音声検出を再度行っているため、重複した音声検出処理が無駄な処理となっている、という問題点を有している。

【0010】また、端末側で入力された音声データの音声検出、分析を行い、分析後のパラメータをサーバへ送信し、サーバ側では送信された分析後のパラメータを用い音声認識を行う構成の従来のサーバ・クライアント型音声認識装置は、認識時に用いるパラメータが決められているため独自仕様のパラメータを用いることができないという問題点や、新規の分析手法を端末に搭載するためのコスト、時間を多く必要とする、という問題点を有している。

【0011】したがって、本発明が解決しようとする主たる課題は、送信時間の短縮、コストの低減を図る音声認識装置及び方法を提供することにある。

【0012】

【課題を解決するための手段】上記課題を解決するための手段を提供する本発明は、その第1のアスペクトにおいて、入力されたデータの音声区間を検出する音声検出部と、前記音声検出部で検出された音声区間の波形データを圧縮する波形圧縮部と、前記波形圧縮部で圧縮された波形データを出力する波形送信部とを有して構成される端末側(クライアント側)の装置と、前記端末側の装置から出力された波形データを受信する波形受信部と、前記波形受信部で受信された波形データを伸張する波形伸張部と、前記波形伸張部で伸張された波形データを用いて認識処理を行い、認識した結果を出力する認識部とを有して構成されるサーバ側の装置とを備えている。

【0013】本発明は、第2のアスペクトにおいて、前記波形圧縮部で圧縮された波形データを一時記憶しながらその波形データを波形送信部へ送信すると共に、入力された波形データとサーバ側から送信された波形データ再送要求信号を受信する波形・信号受信部においてサーバ側からの波形データ再送要求信号が受信された場合に、記憶された波形データを波形送信部へ送信する波形記憶部を有する構成としてもよい。

【0014】本発明は、第3のアスペクトにおいて、サーバ側装置において、端末側装置から出力された波形デ

ータと始端キャンセル信号を受信する波形・信号受信部を備え、認識部では、前記波形・信号受信部で始端キャンセル信号を受信した場合に認識処理を中止する構成としてもよい。

【0015】本発明は、第4のアスペクトにおいて、合成音を合成する音声合成部と、前記音声合成部で合成された合成音の情報を作成し、合成音を出力する合成音情報作成部と、前記音声検出部で検出された音声区間の波形データを圧縮し、前記合成音情報作成部で作成された合成音情報を前記波形データの一部に組み込む波形圧縮部とを有する端末側の装置と、前記波形伸張部で伸張されたデータから合成音情報を取得する合成音情報取得部と、認識が終了した場合に前記合成音情報取得部で取得された合成音情報より合成音と認識結果との対応付けを行い、対応付けされた認識結果、又は認識結果と合成音情報を出力する認識部とを有するサーバ側の装置とを備えている。

【0016】本発明は、第5のアスペクトにおいて、入力された波形データと、サーバ側から送信された波形データ再送要求信号とサーバ側から送信されたサーバ側で利用可能な圧縮方式情報を受信する波形・信号・圧縮方式受信部と、前記波形・信号・圧縮方式受信部においてサーバ側で利用可能な圧縮方式情報を受信した場合に、前記波形・信号・圧縮方式受信部から送られた圧縮方式情報より最適な圧縮方式を選択する圧縮方式選択部と、前記圧縮方式選択部で選択された圧縮方式のインデックスを作成する圧縮方式インデックス作成部と、前記音声検出部で検出された音声区間の波形データを圧縮し、前記合成音情報作成部で作成された合成音情報と、前記圧縮方式インデックス作成部で作成された圧縮方式インデックスを波形データの一部に組み込む波形圧縮部とを有する端末側装置と、前記端末側の装置から送信された波形データと、圧縮方式要求信号を受信する波形・信号受信部と、サーバ側で利用可能な圧縮方式を記憶する圧縮方式記憶部と、前記波形・信号受信部から圧縮方式要求信号が送られた場合に前記圧縮方式記憶部に記憶されている圧縮方式を取得し、端末側装置に圧縮方式情報を出力する圧縮方式取得部と、前記波形伸張部で伸張されたデータから圧縮方式のインデックスを取得する圧縮方式インデックス取得部と、前記圧縮方式インデックス取得部で取得された圧縮方式のインデックスから認識エンジンを選択する認識エンジン選択部と、前記認識エンジン選択部で選択されたエンジンを記憶されているエンジンの中から設定する認識エンジン設定部とを有するサーバ側の装置とを備えている。

【0017】本発明は、第6のアスペクトにおいて、サーバ側装置において、端末側の装置から送信された波形データと始端キャンセル信号と圧縮方式要求信号とコンテンツ側から送信されたタスク情報とを受信する波形・信号・タスク情報受信部と、前記波形・信号・タスク情

10

20

30

40

50

報受信部で受信されたタスク情報を記憶するタスク情報記憶部と、タスク情報とそのタスク使用時に利用可能な 1 個以上の圧縮方式を記憶する圧縮方式・タスク情報対応テーブル記憶部と、前記波形・信号・タスク情報受信部において圧縮方式要求信号を受信した場合に、前記タスク情報記憶部から送られたタスク情報と前記圧縮方式・タスク情報対応テーブル記憶部から送られたタスク情報と圧縮方式との対応テーブルより利用可能な圧縮方式情報を取得し端末側装置に出力する圧縮方式取得部とを備えている。

【0018】本発明は、第 7 のアスペクトにおいて、端末側装置において、入力された波形データと、コンテンツ側から送信されたタスク情報とサーバ側から送信された波形データ再送要求信号とサーバ側から送信されたサーバ側で利用可能な圧縮方式情報とを受信する波形・信号・圧縮方式・タスク情報受信部と、前記波形・信号・圧縮方式・タスク情報受信部で受信されたタスク情報を記憶するタスク情報記憶部と、タスク情報とそのタスク使用時に利用可能な 1 個以上の圧縮方式の対応テーブルを記憶する圧縮方式・タスク情報対応テーブル記憶部と、前記波形・信号・圧縮方式・タスク情報受信部においてサーバ側で利用可能な圧縮方式情報を受信した場合に、前記タスク情報記憶部から送られたタスク情報と、前記圧縮方式・タスク情報対応テーブル記憶部から送られたタスク情報と圧縮方式との対応テーブルと、前記波形・信号・圧縮方式・タスク情報受信部から送られたサーバ側で利用可能な圧縮方式情報をもとに最適な圧縮方式を選択する圧縮方式選択部とを備えている。

【0019】

【発明の実施の形態】発明の実施の形態について説明する。前述した従来の技術の問題点を解決し、上記した課題を解決する本発明は、サーバと端末よりなるサーバ・クライアント型音声認識装置よりなり、端末側で認識用の音声検出を行い、検出後の波形データを圧縮し、圧縮後の波形データをサーバ側に送信し、サーバ側で音声伸張、分析、認識を行うものである。

【0020】〔第 1 の実施の形態〕本発明の第 1 の実施の形態に係る音声認識装置は、図 1 を参照すると、端末（クライアント端末）は、入力されたデータの音声区間を検出する音声検出部（10）と、音声検出部（10）で検出された音声区間の波形データを圧縮する波形圧縮部（20）と、波形圧縮部（20）で圧縮された波形データを出力する波形送信部（30）と、を備え、サーバ側が、端末から出力された波形データを受信する波形受信部（110）と、波形受信部（110）で受信された波形データを伸張する波形伸張部（120）と、波形伸張部（120）で伸張された波形データを用いて音声の分析を行う分析部（130）と、分析部（130）での分析結果（特徴量）と標準パターン記憶部（150）に基づき認識処理を行い、認識した結果を出力する認識部

（140）と、を備えている。この実施の形態では、音声検出後の圧縮された波形データをサーバに送信し、サーバ側で音声分析、認識を行うため、分析後のパラメータ（特徴ベクトル）の仕様を自由に設定することが可能である。

【0021】また、分析部（130）、認識部（140）をサーバ側に備えているため、新規の分析手法、認識手法を搭載するためのコストと時間が少なくて済むという利点を持つ。

10 【0022】また、送信する波形データは圧縮後のデータであるため、送信時に必要とする時間、コストが少なくて済むという利点も併せ持つ。

【0023】上記第 1 の実施の形態に対応した、音声認識方法は、端末側装置において、入力されたデータの音声区間を検出するステップと、前記検出された音声区間の波形データを圧縮するステップと、前記圧縮された波形データをサーバに送信するステップと、を有し、前記サーバ側装置において、前記端末から出力された波形データを受信するステップと、前記受信された波形データを伸張するステップと、前記伸張された波形データを用いて認識処理を行い、認識した結果を出力するステップと、を有する。

20 【0024】〔第 2 の実施の形態〕本発明の第 2 の実施の形態に係るサーバ・クライアント型音声認識装置は、前記第 1 の実施の形態の構成に加え、サーバ側装置において圧縮波形データの受信に失敗した場合に、サーバ側から端末側へ波形データ再送要求信号を送信することで、端末側からサーバ側へ圧縮波形データが再送される仕組み、機構を備えたものである。かかる構成によつて、電波状態が良くない環境においても端末側からサーバ側へ波形データを送信できる割合が高くなる。

30 【0025】より詳細には、本発明の第 2 の実施の形態において、端末側の装置は、図 2 を参照すると、入力された音声の波形データと、前記サーバ側の装置から送信された波形データ再送要求信号を受信する波形・信号受信部（210）を備えている。端末側の装置は、さらに、波形・信号受信部（210）で受信された波形データの音声区間を検出する音声検出部（220）と、音声検出部（220）で検出された音声区間の波形データを圧縮する波形圧縮部（230）と、波形データを出力する波形送信部（240）と、波形圧縮部（230）で圧縮された波形データを一時記憶しながら前記波形データを、前記波形送信部へ送信すると共に、前記波形・信号受信部（210）において、前記サーバ側の装置からの波形データ再送要求信号が受信された場合に、記憶された波形データを、波形送信部（240）へ送信する波形記憶部（250）と、を備えている。

40 【0026】本発明の第 2 の実施の形態において、サーバ側の装置は、端末側の装置から出力された波形データを受信する波形受信部（310）と、波形受信部（31

0) で受信された波形データを伸張する波形伸張部 (320) と、波形伸張部 (320) で伸張された波形データを用いて音声の分析を行う分析部 (340) と、分析部での分析結果に基づき、音声の認識処理を行い、認識した結果を出力する認識部 (350) と、波形受信部 (310) で圧縮波形データの受信に失敗した場合には、波形データの再送要求信号を前記波形受信部から受け取り波形データ再送要求信号を、端末に出力する波形データ再送要求信号送信部 (330) と、を備えている。

【0027】上記第2の実施の形態に対応した、音声認識方法は、端末側装置において前記サーバ側装置から送信される波形データ再送要求信号を受信した場合に、波形記憶部に記憶された波形データを、前記サーバ側装置に送信するステップを有し、前記サーバ装置において、前記端末側装置から送信される圧縮波形データの受信に失敗した場合には、波形データ再送要求信号を前記端末に出力するステップを有する。

【0028】〔第3の実施の形態〕本発明の第3の実施の形態に係るサーバ・クライアント型音声認識装置は、第2の実施の形態の構成に加え、端末側では音声区間を検出しサーバ側へ圧縮した波形データを送信した後音声の始端がキャンセルされた場合に、サーバ側へ始端キャンセル信号を送信する枠組みと、サーバ側では始端キャンセル信号を受信した場合に認識処理を停止することができる枠組みを備えている。かかる構成によって、始端キャンセル時に即座に認識処理を停止することができるため、無駄なメモリ量、計算量を省くことが可能となる。

【0029】より詳細には、本発明の第3の実施の形態に係るサーバ・クライアント型音声認識装置は、図3を参照すると、端末側の装置が、入力された音声の波形データと、前記サーバ側の装置から送信された波形データ再送要求信号を受信する波形・信号受信部 (410) と、波形・信号受信部 (410) で受信された音声の波形データの音声区間を検出する音声検出部 (420) と、音声検出部 (420) で検出された音声区間の波形データを圧縮する波形圧縮部 (430) と、波形データを出力する波形送信部 (440) と、波形圧縮部 (430) で圧縮された波形データを一時記憶しながら前記波形データを、波形送信部 (440) へ送信すると共に、波形・信号受信部 (410) において、前記サーバ側の装置からの波形データ再送要求信号が受信された場合に、記憶された波形データを、前記波形送信部へ送信する波形記憶部 (450) と、音声検出部 (420) で音声の検出後に該検出がキャンセルされた場合、音声検出部 (420) から送信された始端キャンセル時の信号を受け取り始端キャンセル信号を出力する始端キャンセル信号送信部 (460) と、を備えている。

【0030】本発明の第3の実施の形態において、サー

バ側の装置は、端末側の装置から出力された波形データと前記始端キャンセル信号を受信する波形・信号受信部 (510) と、波形・信号受信部 (510) で受信された音声の波形データを伸張する波形伸張部 (520) と、前記波形伸張部で伸張された波形データを用いて音声の分析を行う分析部 (540) と、分析部での分析結果に基づき認識処理を行い、認識した結果を出力する認識部 (550) と、波形・信号受信部 (510) で圧縮波形データの受信に失敗した場合には、波形データの再送要求信号を前記波形受信部から受け取り波形データ再送要求信号を出力する波形データ再送要求信号送信部 (530) と、を備え、認識部 (550) は、前記波形・信号受信部 (510) において始端キャンセル信号を受信した場合に、認識処理を中止する。

【0031】上記第2の実施の形態に対応した、音声認識方法は、端末側装置において、音声の検出後に該検出がキャンセルされた場合、始端キャンセル信号を前記サーバ側装置に送信するステップを有し、サーバ側装置において、前記端末からの前記始端キャンセル信号を受信した場合に、認識処理を中止するステップを有する。

【0032】〔第4の実施の形態〕本発明の第4の実施の形態に係るサーバ・クライアント型音声認識装置は、第3の実施の形態の構成に加え、端末側では合成音を作成、出力し、合成音情報を波形データの一部として組み込み、作成したデータをサーバ側へ送信する枠組みと、サーバ側では端末側から送信されたデータを波形データとその他の情報に分離し、その他の情報から合成音情報を取得し、波形データから得られた認識結果と合成音との対応付けを行う枠組みを備えている。

【0033】より詳細には、本発明の第4の実施の形態において、端末側の装置は、図4を参照すると、合成音を合成する音声合成部 (670) と、音声合成部 (670) で合成された合成音の情報を作成し、合成音を出力する合成音情報作成部 (680) と、を備え、波形圧縮部 (630) は、音声検出部 (620) で検出された音声区間の波形データを圧縮し、合成音情報作成部 (680) で作成された合成音情報を前記波形データの一部に組み込む。

【0034】本発明の第4の実施の形態において、サーバ側の装置は、波形伸張部 (720) で伸張されたデータから合成音情報を取得する合成音情報取得部 (770) を備え、前記認識部 (750) が、認識が終了した場合に前記合成音情報取得部 (770) で取得された合成音情報より合成音と認識結果との対応付けを行い、対応付けされた認識結果、又は認識結果と合成音情報を出力する。

【0035】本発明の第4の実施の形態において、合成音情報については、合成音の内容、合成音出力時のタイムスタンプ情報、合成音出力時の対話の状態番号等、様々な情報を用いることができる。

10

20

30

40

50

【0036】合成音情報を波形データの一部として組み込む手法については、

- ・合成音情報を波形データのヘッダーに含めて圧縮する手法や、
- ・圧縮済みの波形データに合成音情報を付加する手法等を用いることができる。

【0037】合成音情報がサーバ側に送られない場合、端末側からサーバ側へ波形データ送信に失敗した場合に、次に送信に成功した波形データがどの合成音に対して発声された音声であるか判断がつかない、という問題点がある。

【0038】この問題は、この第4の実施の形態によって解決され、合成音情報と波形データが1対1対応となることにより、常に認識結果と合成音の対応付けをとることが可能となり、状態遷移に誤りのない対話システムの構築が可能となる。

【0039】上記第4の実施の形態に対応した、音声認識方法は、端末において、合成音を合成するステップと、前記合成された合成音の情報を作成し、合成音を出力するステップと、検出された音声区間の波形データを圧縮し、前記合成音情報作成部で作成された合成音情報を前記波形データの一部に組み込んで、前記サーバ装置に送信するステップと、を含み、前記サーバ装置において、前記伸張されたデータから合成音情報を取得するステップを有し、認識が終了した場合に、前記取得された合成音情報より合成音と認識結果との対応付けを行い、対応付けされた認識結果、又は認識結果と合成音情報を出力する。

【0040】〔第5の実施の形態〕本発明の第5の実施の形態に係るサーバ・クライアント型音声認識装置は、第4の実施の形態の構成に加え、端末側からサーバ側へ圧縮方式を要求する信号を送信し、サーバ側で圧縮方式要求信号を受信した場合、サーバ側で利用可能な圧縮方式情報を端末側へ送信し、端末側ではサーバ側から送信された圧縮情報と、端末側で利用可能な圧縮方式をもとに最適な圧縮方式を選択する枠組みと、端末側で選択された圧縮情報のインデックスを合成音情報とともに波形データの一部に組み込み、作成したデータをサーバ側へ送信し、サーバ側ではデータを波形データとその他のデータとに分離し、分離後のその他のデータを合成音情報と圧縮方式のインデックスに分離し、圧縮方式のインデックス情報よりその圧縮方式に最適な認識エンジンを選択し、選択された認識エンジンの分析手法、認識手法、標準パターンを用いて認識結果を出力する枠組みを備えている。

【0041】より詳細には、本発明の第5の実施の形態に係るサーバ・クライアント型音声認識装置において、図5を参照すると、端末側の装置は、入力された波形データと、サーバ側の装置から送信された波形データ再送要求信号と、前記サーバ側の装置から送信された前記サ

ーバ側で利用可能な圧縮方式情報を受信する波形・信号・圧縮方式受信部(1010)と、波形・信号・圧縮方式受信部(1010)で受信された波形データの音声区間を検出する音声検出部(1020)と、前記音声検出部で検出された音声区間の波形データを圧縮する波形圧縮部(1030)と、波形データを出力する波形送信部(1040)と、波形圧縮部(1030)で圧縮された波形データを一時記憶しながら前記波形データを、波形送信部(1040)へ送信すると共に、波形・信号・圧縮方式受信部(1010)においてサーバ側の装置からの波形データ再送要求信号を受信された場合に、記憶された波形データを、波形送信部(1040)へ送信する波形記憶部(1050)と、を備えている。

【0042】波形・信号・圧縮方式受信部(1010)において、前記サーバ側で利用可能な圧縮方式情報を受信した場合に、前記波形・信号・圧縮方式受信部(1010)から送られた圧縮方式情報より最適な圧縮方式を選択する圧縮方式選択部(1100)と、圧縮方式選択部(1100)で選択された圧縮方式のインデックスを作成する圧縮方式インデックス作成部(1090)と、を備え、波形圧縮部(1030)は、音声検出部(1020)で検出された音声区間の波形データを圧縮し、合成音情報作成部(1080)で作成された合成音情報と、圧縮方式インデックス作成部(1090)で作成された圧縮方式インデックスを波形データの一部に組み込む。

【0043】本発明の第5の実施の形態において、前記サーバ側の装置は、前記端末側の装置から送信された波形データと、圧縮方式要求信号を受信する波形・信号受信部(1210)と、波形・信号受信部(1210)で受信された音声の波形データを伸張する波形伸張部(1220)と、波形伸張部(1220)で伸張された波形データを用いて音声の分析を行う分析部(1240)と、分析結果に基づき認識処理を行い、認識した結果を出力する認識部(1250)と、前記波形・信号受信部(1210)で圧縮波形データの受信に失敗した場合には、波形データの再送要求信号を前記波形受信部から受け取り波形データ再送要求信号を出力する波形データ再送要求信号送信部(1230)と、前記サーバ側の装置で利用可能な圧縮方式の情報を記憶する圧縮方式記憶部(1320)と、波形・信号受信部(1210)から圧縮方式要求信号が送られた場合に前記圧縮方式記憶部に記憶されている圧縮方式情報を取得し、前記端末側の装置に圧縮方式情報を出力する圧縮方式取得部(1310)と、波形伸張部(1220)で伸張されたデータから圧縮方式のインデックスを取得する圧縮方式インデックス取得部(1280)と、圧縮方式インデックス取得部(1280)で取得された圧縮方式のインデックスから認識エンジンを選択する認識エンジン選択部(1290)と、認識エンジン選択部(1290)で選択された

エンジンを記憶されているエンジンの中から設定する認識エンジン設定部(1300)と、を備える。

【0044】サーバ側、端末側の圧縮方式記憶部(1320)に記憶された圧縮方式については、帯域情報(8kHz、4kHz等)や圧縮方法(フルレート、ハーフレート等)等の情報を含む。また、波形データの一部として合成音情報と圧縮方式のインデックスを含める手法については、合成音情報、圧縮方式のインデックスを波形データのヘッダーに含めて圧縮する手法や、圧縮済みの波形データに合成音情報、圧縮方式のインデックスを付加する手法等を用いることができる。また、端末側において、サーバ側で利用可能な圧縮方式と端末側で利用可能な圧縮方式より最適な圧縮方式を選択する手段については、情報量(データ量)の多い圧縮方式を選択する。例えば、端末側、サーバ側ともにフルレート、ハーフレートの圧縮方式が利用可能な場合、データの情報量の多いフルレートが自動的に選択される。本発明では、選択された圧縮方式によってサーバ側で認識エンジンを選択することが可能であるため、サーバ側で圧縮方式に適した認識エンジンを用意することで、様々な圧縮方式の音声データに対して高い認識率を得ることが可能となる。

【0045】上記第5の実施の形態に対応した、音声認識方法は、端末側装置において、サーバ側装置から送信された前記サーバ側で利用可能な圧縮方式情報を受信するステップと、前記受信した圧縮方式情報により最適な圧縮方式を選択するステップと、前記選択された圧縮方式のインデックスを作成するステップと、音声区間の波形データを圧縮し、作成された圧縮方式インデックスを波形データの一部に組み込んで、前記サーバ装置に送信するステップとを有し、前記サーバ側装置において、前記端末側装置から送信される圧縮方式要求信号を受信した場合に、前記サーバ側で利用可能な圧縮方式の情報を記憶する圧縮方式記憶部に記憶されている圧縮方式情報を取得し、前記端末側装置に圧縮方式情報を出力するステップと、伸張されたデータから圧縮方式のインデックスを取得するステップと、前記取得された圧縮方式のインデックスから認識エンジンを選択するステップと、前記選択されたエンジンを記憶されているエンジンの中から設定するステップと、を有する。

【0046】〔第6の実施の形態〕本発明の第6の実施の形態に係るサーバ・クライアント型音声認識装置は、第5の実施の形態の構成に加え、サーバ側で端末側から送信された圧縮方式要求信号を受信した場合、コンテンツ側より入力されたタスク情報と予めサーバ側で持つタスク情報とそのタスクで使用可能な圧縮方式との対応テーブルより、サーバ側で利用可能な圧縮方式を取得し端末側に送信する枠組みを備える。

【0047】より詳細には、本発明の第6の実施の形態において、サーバ側装置は、図6を参照すると、前記した第5の実施の形態の波形・信号受信部(図5の121

0)に代わって、端末側の装置から送信された波形データと、始端キャンセル信号と、圧縮方式要求信号と、コンテンツ側から送信されたタスク情報と、を受信する波形・信号・タスク情報受信部(1610)を備え、波形・信号・タスク情報受信部(1610)で受信されたタスク情報を記憶するタスク情報記憶部(1730)と、タスク情報と該タスク使用時に利用可能な1個以上の圧縮方式を記憶する圧縮方式・タスク情報対応テーブル記憶部(1720)と、前記波形・信号・タスク情報受信部(1610)において圧縮方式要求信号を受信した場合に、前記タスク情報記憶部(1730)から送られたタスク情報と前記圧縮方式・タスク情報対応テーブル記憶部(1720)から送られたタスク情報と圧縮方式との対応テーブルより利用可能な圧縮方式情報を取得し、前記端末側の装置に出力する圧縮方式取得部(1710)を備えている。

【0048】サーバ側で持つタスク情報とそのタスクで使用可能な圧縮方式との対応テーブル(1720)に記述された圧縮方式については、帯域情報(8kHz、4kHz等)や圧縮方法(フルレート、ハーフレート等)等の情報を含む。

【0049】この第6の実施の形態では、圧縮方式を選択する時にサーバ側と端末側とで共通に利用可能である圧縮方式情報の他に、タスクの情報を加えることによって、サーバ側と端末側とで共通に利用可能な圧縮方式であり、且つ、入力されたタスクに適した圧縮方式を選択することが可能となる。例えば、端末側とサーバ側とで対話を行うタスクでは、レスポンス時間が短いことが望ましいためデータの転送時間の短い圧縮方式をタスク情報と圧縮方式との対応テーブルに記述し、又、端末側で音声入力したキーワードをサーバ側でキーワード検索するようなレスポンス時間を比較的問われないタスクではデータの情報量の多い圧縮方式を、前記テーブルに記述することで、タスクに適した圧縮方式の選択が可能となる。

【0050】上記第6の実施の形態に対応した、音声認識方法は、サーバ側装置において、コンテンツ側から送信されたタスク情報を受信してタスク情報記憶部に記憶するステップと、圧縮方式要求信号を受信した場合に、タスク情報と、圧縮方式との対応テーブルより、利用可能な圧縮方式情報を取得し、前記端末側装置に送信するステップを有する。

【0051】〔第7の実施の形態〕本発明の第7の実施の形態に係るサーバ・クライアント型音声認識装置は、第5の実施の形態の構成に加え、端末側において、サーバ側からサーバ側で利用可能な圧縮方式情報が送信された場合に、サーバ側で利用可能な圧縮方式情報と、コンテンツ側から入力されたタスク情報と、端末側で予め持つタスク情報とそのタスクで使用可能な圧縮方式との対応テーブルより、最適な圧縮方式を選択する枠組みを備

える。

【0052】より詳細には、本発明の第7の実施の形態に係るサーバ・クライアント型音声認識装置は、図7を参照すると、端末側の装置において、前記第5の実施の形態の波形・信号・圧縮方式受信部（図5の1010）に代わって、入力された波形データと、コンテンツ側から送信されたタスク情報と、前記サーバ側の装置から送信された波形データ再送要求信号と、前記サーバ側の装置から送信された前記サーバ側の装置で利用可能な圧縮方式情報とを受信する波形・信号・圧縮方式・タスク情報受信部（1810）を備え、波形・信号・圧縮方式・タスク情報受信部で受信されたタスク情報を記憶するタスク情報記憶部（1930）と、タスク情報とそのタスク使用時に利用可能な1個以上の圧縮方式の対応テーブルを記憶する圧縮方式・タスク情報対応テーブル記憶部（1910）と、前記波形・信号・圧縮方式・タスク情報受信部（1810）においてサーバ側で利用可能な圧縮方式情報を受信した場合に、前記タスク情報記憶部（1930）から送られたタスク情報と、前記圧縮方式・タスク対応テーブル記憶部（1910）から送られたタスク情報と圧縮方式との対応テーブルと、前記波形・信号・圧縮方式・タスク情報受信部（1810）から送られた前記サーバ側の装置で利用可能な圧縮方式情報をもとに最適な圧縮方式を選択する圧縮方式選択部（1900）と、を備える。

【0053】タスク情報とそのタスクで使用可能な圧縮方式との対応テーブルに記述された圧縮方式については、帯域情報(8kHz、4kHz等)や圧縮方法（フルレート、ハーフレート等）等の情報を含む。

【0054】この第7の実施の形態では、前記第6の実施の形態と同様、圧縮方式を選択する時にサーバ側と端末側とで共通に利用可能である圧縮方式情報の他にタスクの情報を加味することによって、サーバ側と端末側とで共通に利用可能な圧縮方式であり、且つ、入力されたタスクに適した圧縮方式を選択することが可能となる。

【0055】上記第7の実施の形態に対応した、音声認識方法は、端末側装置において、コンテンツ側から送信されたタスク情報と、前記サーバ側装置から送信された前記サーバ側で利用可能な圧縮方式情報とを受信するステップと、前記サーバ側で利用可能な圧縮方式情報を受信した場合に、前記タスク情報と、圧縮方式・タスク対応テーブル記憶部から送られたタスク情報と圧縮方式との対応テーブルと、サーバ側の装置で利用可能な圧縮方式情報をもとに最適な圧縮方式を選択するステップと、を有する。

【0056】

【実施例】本発明の実施例について図面を参照して以下に説明する。

【0057】〔実施例1〕図1は、本発明の第1の実施例をなすサーバ・クライアント型音声認識装置の構成を

示す図であり、図1(a)には、端末、図1(b)にはサーバ側の装置構成が示されている。図1(a)を参照すると、端末は、音声検出部10と、波形圧縮部20と、波形送信部30を備えている。端末に入力されたデータは、音声検出部10に送られ音声の検出が行われる。音声の検出方法としては、ワードスポッティング手法等が用いられる。ワードスポッティング手法については、文献2（中川，“確率モデルによる音声認識,”電子情報通信学会,1988）の記載が参照される。検出された音声データは、波形圧縮部20に送られる。波形圧縮部20では、音声検出部10から送られた波形データを圧縮する。圧縮方式としては、VSELP(Vector Sum Excited Linear Prediction)方式、PSI-CELP(Pitch Synchronous Innovation CELP)方式等が用いられる。圧縮された波形データは波形送信部30に送られる。波形送信部30では、波形圧縮部から送られた圧縮波形データを出力する。

【0058】図1(b)を参照すると、サーバ側装置は、波形受信部110と、波形伸張部120と、分析部130と、認識部140と、標準パターン記憶部150とを備えている。端末側の波形送信部30から出力された圧縮波形データを波形受信部110で受信する。波形受信部110では、受信した圧縮波形データを波形伸張部120に送信する。波形伸張部120では、波形受信部110から送られた圧縮波形データを伸張する。伸張された波形データは分析部130に送られる。分析部130では、波形伸張部120で伸張された波形データを特徴ベクトルに変換する。特徴ベクトルとしてはケプストラム、Δケプストラム等が用いられる。Δケプストラムは、文献3（古井，“デジタル音声処理,”東海大学出版会,pp44-47,1985）等の記載が参照される。分析された特徴ベクトルは認識部140に送られる。認識部140では、分析部130から送られた特徴ベクトルと、標準パターン記憶部150から送られた標準パターンとの間の尤度を計算し、認識結果を求める。尤度の計算方法は、Viterbiアルゴリズム、DPマッチング等が用いられる。求められた認識結果は出力される。Viterbiアルゴリズムについては、文献2（中川，“確率モデルによる音声認識,”電子情報通信学会,1988）、DPマッチングについては、文献4（迫江，“傾斜制限DPマッチングによる音声認識,”音声講論,pp67-68,1974）の記載が参照される。

【0059】〔実施例2〕次に、本発明の第2の実施例について説明する。図2は、本発明の第2の実施例をなすサーバ・クライアント型音声認識装置の構成を示す図であり、図2(a)には端末、図2(b)にはサーバ側の装置構成が示されている。図2(a)を参照すると、端末は、波形・信号受信部210と、音声検出部220と、波形圧縮部230と、波形送信部240、波形記憶部250を備えている。端末の波形・信号受信部210

には、音声データ、又はサーバから送信された波形データ再送要求信号が入力される。音声データが入力された場合、該音声データは、音声検出部220に送られる。音声検出部220では、波形・信号受信部210から送られた波形データの音声検出を行う。検出方法としてはワードスポッティング手法等が用いられる。検出された音声データは波形圧縮部230に送られる。波形圧縮部230では、音声検出部220から送られた波形データを圧縮する。圧縮方式としては、VSELP方式、PSI-CELP方式等が用いられる。圧縮された波形データは波形記憶部250に送られる。波形記憶部250では、波形圧縮部230から送られた圧縮波形データを波形送信部240に送信し、同一データを記憶する。波形送信部240では、波形記憶部250から送られた圧縮波形データを出力する。

【0060】波形・信号受信部210に、サーバから送信された波形データ再送要求信号が入力された場合、受信された信号は、波形記憶部250に送られる。波形記憶部250では記憶されている圧縮波形データを波形送信部240に送信する。波形送信部240では、波形記憶部250から送られた圧縮波形データを出力する。

【0061】図2(b)を参照すると、サーバ側は、波形受信部310と、波形伸張部320と、波形データ再送要求信号送信部330と、分析部340と、認識部350と、標準パターン記憶部360とを備えている。サーバ側は、端末側の波形送信部240から出力された圧縮波形データを波形受信部310で受信する。

【0062】波形受信部310では、受信した圧縮波形データを波形伸張部320に送信する。圧縮波形データの受信に失敗した場合は、波形データ再送要求信号送信部330に波形データの再送要求信号を送信する。波形データ再送要求信号送信部330では、波形受信部310から送られた信号を出力する。波形伸張部320では、波形受信部310から送られた圧縮波形データを伸張する。伸張された波形データは分析部340に送られる。分析部340では、波形伸張部320で伸張された波形データを特徴ベクトルに変換する。特徴ベクトルとしてはケプストラム、 Δ ケプストラム等が用いられる。分析された特徴ベクトルは認識部350に送られる。認識部350では、分析部340から送られた特徴ベクトルと、標準パターン記憶部360から送られた標準パターンとの間の尤度を計算し、認識結果を求める。尤度の計算方法はViterbiアルゴリズム、DPマッチング等が用いられる。求められた認識結果は出力される。

【0063】〔実施例3〕次に、本発明の第3の実施例について説明する。図3は、本発明の第3の実施例をなすサーバ・クライアント型音声認識装置の構成を示す図であり、図3(a)には端末、図3(b)にはサーバ側の装置構成が示されている。図3(a)を参照すると、端末は、波形・信号受信部410、音声検出部420

と、波形圧縮部430と、波形送信部440、波形記憶部450、始端キャンセル信号送信部460を備えている。端末の波形・信号受信部410には音声データ、又はサーバから送信された波形データ再送要求信号が入力される。音声データが入力された場合、データは音声検出部420に送られる。音声検出部420では、波形・信号受信部410から送られた波形データの音声検出を行う。検出方法としてはワードスポッティング手法等が用いられる。検出された音声データは波形圧縮部430に送られる。

【0064】音声の検出後に検出がキャンセルされた場合、始端キャンセル時の信号が始端キャンセル信号送信部460に送信される。始端キャンセル信号送信部460では、送られた始端キャンセル時の信号を出力する。

【0065】波形圧縮部430では、音声検出部420から送られた波形データを圧縮する。圧縮方式としては、VSELP方式、PSI-CELP方式等が用いられる。圧縮された波形データは波形記憶部450に送られる。波形記憶部450では、波形圧縮部430から送られた圧縮波形データを波形送信部440に送信し、同一データを記憶する。波形送信部440では、波形圧縮部430から送られた圧縮波形データを出力する。波形・信号受信部410にサーバから送信された波形データ再送要求信号が入力された場合、受信された信号は波形記憶部450に送られる。波形記憶部450では記憶されている圧縮波形データを波形送信部440に送信する。波形送信部440では波形記憶部450から送られた圧縮波形データを出力する。

【0066】図3(b)を参照すると、サーバ側は、波形・信号受信部510と、波形伸張部520と、波形データ再送要求信号送信部530と、分析部540と、認識部550と、標準パターン記憶部560とを備えている。

【0067】サーバ側では、端末側の波形送信部440から出力された圧縮波形データ、又は、始端キャンセル信号送信部460から出力された始端キャンセル時の信号を波形・信号受信部510で受信する。

【0068】波形・信号受信部510では、受信した圧縮波形データを波形伸張部520に送信する。圧縮波形データの受信に失敗した場合は、波形データ再送要求信号送信部530に波形データ再送要求信号を送信する。波形データ再送要求信号送信部530では、波形・信号受信部510から送られた信号を出力する。波形伸張部520では、波形・信号受信部510から送られた圧縮波形データを伸張する。伸張された波形データは分析部540に送られる。分析部540では、波形伸張部520で伸張された波形データを特徴ベクトルに変換する。特徴ベクトルとしてはケプストラム、 Δ ケプストラム等が用いられる。分析された特徴ベクトルは認識部550に送られる。認識部550では、分析部540から送ら

れた特徴ベクトルと、標準パターン記憶部560から送られた標準パターンとの間の尤度を計算し、認識結果を求める。尤度の計算方法はViterbiアルゴリズム、DPマッチング等が用いられる。求められた認識結果は出力される。波形・信号受信部510に始端キャンセル信号送信部460から送信された始端キャンセル時の信号が入力された場合、信号は認識部550に送られる。

【0069】認識部550では、波形・信号受信部510から始端キャンセル時の信号が送られてきた場合、認識を途中で中断する。

【0070】〔実施例4〕次に、本発明の第4の実施例について説明する。図4は、本発明の第4の実施例をなすサーバ・クライアント型音声認識装置の構成を示す図であり、図4(a)には端末、図4(b)にはサーバ側の装置構成が示されている。図4(a)を参照すると、端末は、波形・信号受信部610と、音声検出部620と、波形圧縮部630と、波形送信部640と、波形記憶部650と、始端キャンセル信号送信部660と、音声合成部670と、合成音情報作成部680と、を備えている。端末の波形・信号受信部610には音声データ、又はサーバから送信された波形データ再送要求信号が入力される。音声データが入力された場合、データは音声検出部620に送られる。音声検出部620では、波形・信号受信部610から送られた波形データの音声検出を行う。検出方法としてはワードスポッティング手法等が用いられる。検出された音声データは波形圧縮部630に送られる。

【0071】音声の検出後に検出がキャンセルされた場合、始端キャンセル時の信号が始端キャンセル信号送信部660に送信される。始端キャンセル信号送信部660では、送られた始端キャンセル時の信号を出力する。

【0072】本実施例では、更に、合成音を作成する音声合成部670を持つ。音声合成部670で作成された合成音は、合成音情報作成部680に送られる。合成音情報作成部680では、合成音情報を作成し、作成した合成音情報を波形圧縮部630に送信し、音声合成部670から送られた合成音を出力する。

【0073】合成音情報については、合成音の内容や合成音出力時のタイムスタンプ情報等が用いられる。

【0074】波形圧縮部630では、音声検出部620から送られた波形データを圧縮し、合成音情報作成部680で取得された合成音情報を波形データの一部に含める。

【0075】波形データの一部に合成音情報を含める手法としては、合成音情報を波形データのヘッダーに含めて圧縮する方法や、圧縮した波形データに合成音情報を付加する方法等が用いられる。

【0076】圧縮方式としては、VSELP方式、PSI-CELP方式等が用いられる。圧縮された波形データは波形記憶部650に送られる。波形記憶部650では、波形圧縮

部630から送られた圧縮波形データを波形送信部640に送信し、同一データを記憶する。波形送信部640では、波形圧縮部630から送られた波形データを出力する。波形・信号受信部610にサーバから送信された波形データ再送要求信号が入力された場合、受信された信号は波形記憶部650に送られる。波形記憶部650では記憶されている圧縮波形データを波形送信部640に送信する。波形送信部640では波形記憶部650から送られた圧縮波形データを出力する。

10 【0077】図4(b)を参照すると、サーバ側では、波形・信号受信部710と、波形伸張部720と、波形データ再送要求信号送信部730と、分析部740と、認識部750と、標準パターン記憶部760と、合成音情報取得部770を備えている。サーバ側では、端末側の波形送信部640から出力された圧縮波形データ、又は始端キャンセル信号送信部660から出力された始端キャンセル時の信号を波形・信号受信部710で受信する。波形・信号受信部710では、受信した圧縮波形データを波形伸張部720に送信する。圧縮波形データの受信に失敗した場合は、波形データ再送要求信号送信部730に波形データ再送要求信号を送信する。波形データ再送要求信号送信部730では、波形・信号受信部710から送られた信号を出力する。

20 【0078】波形伸張部720では、波形・信号受信部710から送られたデータを波形データ合成音情報等のその他の情報とに分離し、波形データを伸張する。その他の情報は合成音情報取得部770に送られ、伸張された波形データは分析部740に送られる。

30 【0079】合成音情報取得部770では、波形伸張部720から送られた情報から合成音情報を取得する。取得された合成音情報は認識部750に送られる。

【0080】分析部740では、波形伸張部720で伸張された波形データを特徴ベクトルに変換する。特徴ベクトルとしてはケプストラム、 Δ ケプストラム等が用いられる。分析された特徴ベクトルは認識部750に送られる。認識部750では、分析部740から送られた特徴ベクトルと、標準パターン記憶部760から送られた標準パターンとの間の尤度を計算し、認識結果を求める。尤度の計算方法はViterbiアルゴリズム、DPマッチング等が用いられる。

【0081】求められた認識結果と合成音情報取得部770で取得された合成音情報は共に端末側に出力されるか、又は、両者が関連付けられることでどの合成音に対する認識結果であるかが対応付けされ、対応付けされた後の認識結果が端末側に出力される。

40 【0082】波形・信号受信部710に始端キャンセル信号送信部660から送信された始端キャンセル時の信号が入力された場合、信号は認識部750に送られる。認識部750では、波形・信号受信部710から始端キャンセル時の信号が送られてきた場合、認識を途中で中

断する。

【0083】〔実施例5〕次に、本発明の第5の実施例について説明する。図5は、本発明の第5の実施例をなすサーバ・クライアント型音声認識装置の構成を示す図であり、図5(a)には端末、図5(b)にはサーバ側の装置構成が示されている。図5(a)を参照すると、端末は、波形・信号・圧縮方式受信部1010と、音声検出部1020と、波形圧縮部1030と、波形送信部1040と、波形記憶部1050と、始端キャンセル信号送信部1060と、音声合成部1070と、合成音情報作成部1080と、圧縮方式インデックス作成部1090と、圧縮方式選択部1100と、圧縮方式記憶部1110と、圧縮方式要求信号送信部1120とを備えている。端末の波形・信号・圧縮方式受信部1010には、音声データ、又は圧縮方式要求信号、又は、サーバから送信された波形データ再送要求信号、又はサーバから送信された圧縮方式情報が入力される。

【0084】波形・信号・圧縮方式受信部1010に圧縮方式要求信号が入力された場合、圧縮方式要求信号は波形・信号・圧縮方式受信部1010から圧縮方式要求信号送信部1120に送られる。

【0085】圧縮方式要求信号送信部1120では、波形・信号・圧縮方式受信部1010から送られた圧縮方式要求信号をサーバ側装置へ出力する。

【0086】波形・信号・圧縮方式受信部1010に、サーバ側から送信された圧縮方式情報が入力された場合、圧縮方式情報は、圧縮方式選択部1100へ送られる。

【0087】本実施例では、端末側で利用可能な圧縮方式を記憶する圧縮方式記憶部1110、波形を圧縮する方式を選択する圧縮方式選択部1100を備えている。

【0088】圧縮方式記憶部1110に記憶された圧縮方式については、波形圧縮方法（フルレート、ハーフレート等）や帯域（4kHz、8kHz等）の情報を含む。

【0089】圧縮方式選択部1100では、圧縮方式記憶部1110に記憶されている圧縮方式情報と、波形・信号・圧縮方式受信部1010から送られたサーバ側で利用可能な圧縮方式をもとに、データの情報量の多い圧縮方式を選択する。例えば、端末側でもサーバ側でも、ともに、フルレートとハーフレートの圧縮方式を持つ場合、データの情報量の多いフルレートが選択される。

【0090】圧縮方式選択部1100で選択された方式名は、圧縮方式インデックス作成部1090に送られる。

【0091】圧縮方式インデックス作成部1090では、圧縮方式選択部で選択された方式名のインデックスを作成する。圧縮方式インデックス作成部1090で作成されたインデックスは波形圧縮部1030に送られる。

【0092】波形・信号・圧縮方式受信部1010に音

声データが入力された場合、データは音声検出部1020に送られる。音声検出部1020では、波形・信号・圧縮方式受信部1010から送られた波形データの音声検出を行う。検出方法としてはワードスポットティング手法等が用いられる。検出された音声データは波形圧縮部1030に送られる。

【0093】音声の検出後に検出がキャンセルされた場合、始端キャンセル時の信号が始端キャンセル信号送信部1060に送信される。始端キャンセル信号送信部1060では、送られた始端キャンセル時の信号を出力する。

【0094】本実施例では、第4の実施例と同様に合成音を作成する音声合成部1070を持つ。音声合成部1070で作成された合成音は合成音情報作成部1080に送られる。

【0095】合成音情報作成部1080では、合成音情報を作成し、作成した合成音情報を波形圧縮部1030に送信し、音声合成部1070から送られた合成音を出力する。合成音情報については、合成音の内容や合成音出力時のタイムスタンプ情報等が用いられる。波形圧縮部1030では、音声検出部1020から送られた波形データを圧縮し、合成音情報作成部1080で作成された合成音情報と、圧縮方式インデックス作成部1090で作成された圧縮方式のインデックスとを前記波形データの一部に含める。

【0096】波形データの一部に合成音情報、圧縮方式のインデックスを含める手法としては、合成音情報、圧縮方式のインデックスを波形データのヘッダーに含めて圧縮する方法や、圧縮した波形データに合成音情報、圧縮方式のインデックスを付加する方法等が用いられる。圧縮方式としては、VSELP方式、PSI-CELP方式等が用いられる。圧縮された波形データは波形記憶部1050に送られる。

【0097】波形記憶部1050では、波形圧縮部1030から送られた圧縮波形データを波形送信部1040に送信し、同一データを記憶する。

【0098】波形送信部1040では、波形圧縮部1030から送られた圧縮波形データをサーバ側装置へ出力する。波形・信号・圧縮方式受信部1010にサーバから送信された波形データ再送要求信号が入力された場合、受信された信号は波形記憶部1050に送られる。波形記憶部1050では記憶されている圧縮波形データを波形送信部1040に送信する。波形送信部1040では波形記憶部1050から送られた圧縮波形データを出力する。

【0099】図5(b)を参照すると、サーバ側では、波形・信号受信部1210と、波形伸張部1220と、波形データ再送要求信号送信部1230と、分析部1240と、認識部1250と、標準パターン記憶部1260と、合成音情報取得部1270と、圧縮方式インデッ

クス取得部 1280 と、認識エンジン選択部 1290 と、認識エンジン設定部 1300 と、圧縮方式取得部 1310 と、圧縮方式記憶部 1320 と、を備えている。

【0100】サーバ側では、端末側の波形送信部 1040 から出力された圧縮波形データ、又は始端キャンセル信号送信部 1060 から出力された始端キャンセル時の信号、又は圧縮方式要求信号送信部 1120 から出力された波形圧縮方式要求信号を波形・信号受信部 1210 で受信する。波形・信号受信部 1210 では、波形圧縮方式要求信号が受信された場合、受信信号を圧縮方式取得部 1310 へ送る。

【0101】圧縮方式取得部 1310 では、波形・信号受信部 1210 から波形圧縮方式要求信号が送られた場合、圧縮方式記憶部 1320 に記憶されている圧縮方式の中からサーバ側で利用可能な圧縮方式を取得し、取得した圧縮方式情報を端末側に出力する。波形・信号受信部 1210 において波形データが受信された場合、波形伸張部 1220 に送信する。

【0102】波形データの受信に失敗した場合は、波形データ再送要求信号送信部 1230 に波形データ再送要求信号を送信する。波形データ再送要求信号送信部 1230 では、波形・信号受信部 1210 から送られた信号を出力する。波形伸張部 1220 では、波形・信号受信部 1210 から送られたデータを波形データとその他のデータとに分離し、波形データを伸張する。

【0103】その他の情報は合成音情報取得部 1270 と圧縮方式インデックス取得部 1280 に送られ、伸張された波形データは分析部 1240 に送られる。合成音情報取得部 1270 では、波形伸張部 1220 から送られた情報から合成音情報を取得する。取得された合成音情報は認識部 1250 に送られる。

【0104】圧縮方式インデックス取得部 1280 では、波形伸張部 1220 から送られた情報から圧縮方式のインデックスを取得する。

【0105】取得されたインデックスは認識エンジン選択部 1290 に送られる。

【0106】認識エンジン選択部 1290 では、圧縮方式インデックス取得部 1280 で取得されたインデックスをもとに認識に用いる認識エンジンを選択する。選択されたエンジンの情報は認識エンジン設定部 1300 と標準パターン記憶部 1260 に送られる。

【0107】認識エンジン設定部 1300 では、認識エンジン選択部 1290 から送られたエンジン情報を基に、使用するエンジンの分析手法を分析部 1240 へ送り、認識手法を認識部 1250 へ送る。

【0108】標準パターン記憶部 1260 では、認識エンジン選択部 1290 から送られたエンジン情報より使用する標準パターンを設定する。

【0109】分析部 1240 では、波形伸張部 1220 で伸張された波形データを特徴ベクトルに変換する。特

徴ベクトルとしてはケプストラム、 Δ ケプストラム等が用いられる。分析された特徴ベクトルは認識部 1250 に送られる。認識部 1250 では、分析部 1240 から送られた特徴ベクトルと、標準パターン記憶部 1260 から送られた標準パターンとの間の尤度を計算し、認識結果を求める。尤度の計算方法は Viterbi アルゴリズム、DP マッチング等が用いられる。求められた認識結果と合成音情報取得部 1270 で取得された合成音情報は共に端末側に出力されるか、又は、両者が関連付けられることでどの合成音に対する認識結果であるかが対応付けされ、対応付けされた後の認識結果が端末側に出力される。波形・信号受信部 1210 に始端キャンセル信号送信部 1060 から送信された始端キャンセル時の信号が入力された場合、信号は認識部 1250 に送られる。認識部 1250 では、波形・信号受信部 1210 から始端キャンセル時の信号が送られてきた場合、認識を途中で中断する。

【0110】【実施例 6】次に、本発明の第 6 の実施例について説明する。図 6 は、本発明の第 6 の実施例をなすサーバ・クライアント型音声認識装置の構成を示す図であり、図 6 (a) には端末、図 6 (b) にはサーバ側の装置構成が示されている。図 6 (a) を参照すると、端末は、波形・信号・圧縮方式受信部 1410 と、音声検出部 1420 と、波形圧縮部 1430 と、波形送信部 1440 と、波形記憶部 1450 と、始端キャンセル信号送信部 1460 と、音声合成部 1470 と、合成音情報作成部 1480 と、圧縮方式インデックス作成部 1490 と、圧縮方式選択部 1500 と、圧縮方式記憶部 1510 と、圧縮方式要求信号送信部 1520 と、を備えている。端末の波形・信号・圧縮方式受信部 1410 には、音声データ、又は圧縮方式要求信号、又はサーバから送信された波形データ再送要求信号、又はサーバから送信された圧縮方式情報が入力される。圧縮方式要求信号が入力された場合、圧縮方式要求信号は圧縮方式要求信号送信部 1520 に送られる。圧縮方式要求信号送信部 1520 では、波形・信号・圧縮方式受信部 1410 から送られた圧縮方式要求信号をサーバ側装置へ出力する。

【0111】波形・信号・圧縮方式受信部 1410 にサーバ側から送信されたサーバ側で利用可能な圧縮方式情報が入力された場合、圧縮方式情報は圧縮方式選択部 1500 へ送られる。

【0112】本実施例では、第 4 の実施例と同様に、端末側で利用可能な圧縮方式を記憶する圧縮方式記憶部 1510、波形を圧縮する方式を選択する圧縮方式選択部 1500 を持つ。

【0113】圧縮方式記憶部 1510 に記憶された圧縮方式については、波形圧縮方法（フルレート、ハーフレート等）や帯域（4kHz、8kHz 等）の情報を含む。圧縮方式選択部 1500 では、圧縮方式記憶部 1510 に記憶

されている端末側で利用可能な圧縮方式情報と、波形・信号・圧縮方式受信部 1410 から送られたサーバ側で利用可能な圧縮方式をもとにデータの情報量の多い圧縮方式を選択する。例えば、端末側でもサーバ側でもともにフルレートとハーフレートの圧縮方式を持つ場合、データの情報量の多いフルレートが選択される。圧縮方式選択部 1500 で選択された方式名は圧縮方式インデックス作成部 1490 に送られる。圧縮方式インデックス作成部 1490 では、圧縮方式選択部で選択された方式名のインデックスを作成する。圧縮方式インデックス作成部 1490 で作成されたインデックスは波形圧縮部 1430 に送られる。

【0114】波形・信号・圧縮方式受信部 1410 に音声データが入力された場合、データは音声検出部 1420 に送られる。音声検出部 1420 では、波形・信号・圧縮方式受信部 1410 から送られた波形データの音声検出を行う。検出方法としてはワードスポッティング手法等が用いられる。検出された音声データは波形圧縮部 1430 に送られる。音声の検出後に検出がキャンセルされた場合、始端キャンセル時の信号が始端キャンセル信号送信部 1460 に送信される。始端キャンセル信号送信部 1460 では、送られた始端キャンセル時の信号をサーバ側装置へ出力する。

【0115】本実施例では第 4、第 5 の実施例と同様に、合成音を作成する音声合成部 1470 を持つ。音声合成部 1470 で作成された合成音は合成音情報作成部 1480 に送られる。

【0116】合成音情報作成部 1480 では、合成音情報を作成し、作成した合成音情報を波形圧縮部 1430 に送信し、音声合成部 1470 から送られた合成音を出力する。合成音情報については、合成音の内容や合成音出力時のタイムスタンプ情報等が用いられる。

【0117】波形圧縮部 1430 では、音声検出部 1420 から送られた波形データを圧縮し、合成音情報作成部 1480 で作成された合成音情報と、圧縮方式インデックス作成部 1490 で作成された圧縮方式のインデックスとを前記波形データの一部に含める。波形データの一部に合成音情報、圧縮方式のインデックスを含める手法としては、

- ・合成音情報、圧縮方式のインデックスを波形データのヘッダーに含めて圧縮する方法や、
- ・圧縮した波形データに合成音情報、圧縮方式のインデックスを付加する方法

等が用いられる。

【0118】圧縮方式としては、VSELP 方式、PSI-CELP 方式等が用いられる。圧縮された波形データは波形記憶部 1450 に送られる。波形記憶部 1450 では、波形圧縮部 1430 から送られた圧縮波形データを波形送信部 1440 に送信し、同一データを記憶する。波形送信部 1440 では、波形圧縮部 1430 から送られた圧縮

波形データをサーバ側装置へ出力する。波形・信号・圧縮方式受信部 1410 にサーバから送信された波形データ再送要求信号が入力された場合、受信された信号は波形記憶部 1450 に送られる。波形記憶部 1450 では記憶されている圧縮波形データを波形送信部 1440 に送信する。波形送信部 1440 では波形記憶部 1450 から送られた圧縮波形データを出力する。

【0119】図 6 (b) を参照すると、サーバ側では、波形・信号・タスク情報受信部 1610 と、波形伸張部 1620 と、波形データ再送要求信号送信部 1630 と、分析部 1640 と、認識部 1650 と、標準パターン記憶部 1660 と、合成音情報取得部 1670 と、圧縮方式インデックス取得部 1680 と、認識エンジン選択部 1690 と、認識エンジン設定部 1700 と、圧縮方式取得部 1710 と、圧縮方式・タスク情報対応テーブル記憶部 1720 と、タスク情報記憶部 1730 とを備えている。

【0120】サーバ側では、端末側の波形送信部 1440 から出力された波形データ、又は始端キャンセル信号送信部 1460 から出力された始端キャンセル時の信号、又は圧縮方式要求信号送信部 1520 から出力された波形圧縮方式要求信号、又はコンテンツ側より入力されたタスク情報を波形・信号・タスク情報受信部 1610 で受信する。

【0121】波形・信号・タスク情報受信部 1610 にコンテンツ側よりタスク情報が入力された場合、タスク情報記憶部 1730 に送信する。タスク情報記憶部 1730 では、波形・信号・タスク情報受信部 1610 より送られたタスク情報を記憶し、同一情報を圧縮方式取得部 1710 に送る。

【0122】波形・信号・タスク情報受信部 1610 において波形圧縮方式要求信号が受信された場合、受信信号を圧縮方式取得部 1710 へ送る。圧縮方式取得部 1710 では、波形・信号・タスク情報受信部 1610 から波形圧縮方式要求信号が送られた場合、圧縮方式・タスク情報対応テーブル記憶部 1720 に記憶されているタスク情報とそのタスクで利用可能な圧縮方式との対応テーブルと、タスク情報記憶部 1730 から送られたタスク情報とから利用可能な圧縮方式を取得し、取得した圧縮方式情報を端末側に出力する。

【0123】波形・信号・タスク情報受信部 1610 において波形データが受信された場合、波形伸張部 1620 に送信する。波形データの受信に失敗した場合は、波形データ再送要求信号送信部 1630 に波形データ再送要求信号を送信する。波形データ再送要求信号送信部 1630 では、波形・信号・タスク情報受信部 1610 から送られた信号を出力する。波形伸張部 1620 では、波形・信号・タスク情報受信部 1610 から送られたデータを波形データとその他のデータとに分離し、波形データを伸張する。

【0124】その他の情報は合成音情報取得部1670と圧縮方式インデックス取得部1680に送られ、伸張された波形データは分析部1640に送られる。合成音情報取得部1670では、波形伸張部1620から送られた情報から合成音情報を取得する。取得された合成音情報は認識部1650に送られる。

【0125】圧縮方式インデックス取得部1680では、波形伸張部1620から送られた情報から圧縮方式のインデックスを取得する。取得されたインデックスは認識エンジン選択部1690に送られる。

【0126】認識エンジン選択部1690では、圧縮方式インデックス取得部1680で取得されたインデックスをもとに認識に用いる認識エンジンを選択する。選択されたエンジンの情報は認識エンジン設定部1700と標準パターン記憶部1660に送られる。

【0127】認識エンジン設定部1700では、認識エンジン選択部1690から送られたエンジン情報を元に使用するエンジンの分析手法を分析部1640へ送り、認識手法を認識部1650へ送る。標準パターン記憶部1660では、認識エンジン選択部1690から送られたエンジン情報より使用する標準パターンを設定する。分析部1640では、波形伸張部1620で伸張された波形データの特徴ベクトルに変換する。特徴ベクトルとしてはケプストラム、Δケプストラム等が用いられる。

【0128】分析された特徴ベクトルは認識部1650に送られる。認識部1650では、分析部1640から送られた特徴ベクトルと、標準パターン記憶部1660から送られた標準パターンとの間の尤度を計算し、認識結果を求める。尤度の計算方法はViterbiアルゴリズム、DPマッチング等が用いられる。求められた認識結果と合成音情報取得部1670で取得された合成音情報は共に端末側に出力されるか、又は、両者が関連付けられることでどの合成音に対する認識結果であるかが対応付けられ、対応付けされた後の認識結果が端末側に出力される。波形・信号・タスク情報受信部1610に始端キャンセル信号送信部1460から送信された始端キャンセル時の信号が入力された場合、信号は認識部1650に送られる。認識部1650では、波形・信号・タスク情報受信部1610から始端キャンセル時の信号が送られてきた場合、認識を途中で中断する。

【0129】〔実施例7〕次に、本発明の第7の実施例について説明する。図7は、本発明の第7の実施例をなすサーバ・クライアント型音声認識装置の構成を示す図であり、図7(a)には端末、図7(b)にはサーバ側の装置構成が示されている。図7(a)を参照すると、端末は、波形・信号・圧縮方式・タスク情報受信部1810と、音声検出部1820と、波形圧縮部1830と、波形送信部1840と、波形記憶部1850と、始端キャンセル信号送信部1860と、音声合成部1870と、合成音情報作成部1880と、圧縮方式インデッ

クス作成部1890と、圧縮方式選択部1900と、圧縮方式・タスク情報対応テーブル記憶部1910と、圧縮方式要求信号送信部1920と、タスク情報記憶部1930と、を備えている。

【0130】図7を参照すると、この実施例において、端末側の波形・信号・圧縮方式・タスク情報受信部1810には、音声データ、又は圧縮方式要求信号、又はサーバから送信された波形データ再送要求信号、又はサーバから送信された圧縮方式情報、又はコンテンツ側から送信されたタスク情報が入力される。

【0131】波形・信号・圧縮方式・タスク情報受信部1810に、タスク情報が入力された場合、入力されたタスク情報はタスク情報記憶部1930に送られる。タスク情報記憶部1930では、波形・信号・圧縮方式・タスク情報受信部1810から送られたタスク情報を圧縮方式選択部1900に送り、記憶する。

【0132】波形・信号・圧縮方式・タスク情報受信部1810に、圧縮方式要求信号が入力された場合、圧縮方式要求信号は圧縮方式要求信号送信部1920に送られる。圧縮方式要求信号送信部1920では、波形・信号・圧縮方式・タスク情報受信部1810から送られた圧縮方式要求信号をサーバ側装置へ出力する。

【0133】波形・信号・圧縮方式・タスク情報受信部1810に、サーバ側から送信されたサーバ側で利用可能な圧縮方式情報が入力された場合、圧縮方式情報は圧縮方式選択部1900へ送られる。

【0134】本実施例では、タスク情報とそのタスクで利用可能な圧縮方式との対応テーブル記憶する圧縮方式・タスク情報対応テーブル記憶部1910を持つ。圧縮方式・タスク情報対応テーブル記憶部に記憶された圧縮方式については、波形圧縮方法（フルレート、ハーフレート等）や帯域（4kHz、8kHz等）の情報を含む。

【0135】圧縮方式選択部1900では、タスク情報記憶部1930から送られたタスク情報と、波形・信号・圧縮方式・タスク情報受信部1810から送られたサーバ側で利用可能な圧縮方式情報と、圧縮方式・タスク情報対応テーブル記憶部1910に記憶されている対応テーブル情報とをもとにタスク情報が加味されたデータの情報量の多い圧縮方式を選択する。圧縮方式選択部1900で選択された方式名は圧縮方式インデックス作成部1890に送られる。

【0136】圧縮方式インデックス作成部1890では、圧縮方式選択部で選択された方式名のインデックスを作成する。圧縮方式インデックス作成部1890で作成されたインデックスは波形圧縮部1830に送られる。

【0137】波形・信号・圧縮方式・タスク情報受信部1810に音声データが入力された場合、データは音声検出部1820に送られる。音声検出部1820では、波形・信号・圧縮方式・タスク情報受信部1810から

10

20

30

40

50

送られた波形データの音声検出を行う。検出方法としてはワードスポッティング手法等が用いられる。検出された音声データは波形圧縮部 1830 に送られる。音声の検出後に検出がキャンセルされた場合、始端キャンセル時の信号が始端キャンセル信号送信部 1860 に送信される。始端キャンセル信号送信部 1860 では、送られた始端キャンセル時の信号を出力する。

【0138】本実施例では、第 4、第 5、第 6 の実施例と同様に、合成音を作成する音声合成部 1870 を持つ。音声合成部 1870 で作成された合成音は合成音情報作成部 1880 に送られる。合成音情報作成部 1880 では、合成音情報を作成し、作成した合成音情報を波形圧縮部 1830 に送信し、音声合成部 1870 から送られた合成音を出力する。合成音情報については、合成音の内容や合成音出力時のタイムスタンプ情報等が用いられる。

【0139】波形圧縮部 1830 では、音声検出部 1820 から送られた波形データを圧縮し、合成音情報作成部 1880 で作成された合成音情報と、圧縮方式インデックス作成部 1890 で作成された圧縮方式のインデックスとを前記波形データの一部に含める。波形データの一部に合成音情報、圧縮方式のインデックスを含める手法としては、・合成音情報、圧縮方式のインデックスを波形データのヘッダーに含めて圧縮する方法や、・圧縮した波形データに合成音情報、圧縮方式のインデックスを付加する方法等が用いられる。圧縮方式としては、VSELP 方式、PSI-CELP 方式等が用いられる。圧縮された波形データは波形記憶部 1850 に送られる。波形記憶部 1850 では、波形圧縮部 1830 から送られた圧縮波形データを波形送信部 1840 に送信し、同一データを記憶する。波形送信部 1840 では、波形圧縮部 1830 から送られた圧縮波形データをサーバ側装置へ出力する。

【0140】波形・信号・圧縮方式・タスク情報受信部 1810 に、サーバから送信された波形データ再送要求信号が入力された場合、受信された信号は波形記憶部 1850 に送られる。波形記憶部 1850 では記憶されている圧縮波形データを波形送信部 1840 に送信する。波形送信部 1840 では波形記憶部 1850 から送られた圧縮波形データを出力する。

【0141】図 7 (b) を参照すると、サーバ側では、波形・信号受信部 2010 と、波形伸張部 2020 と、波形データ再送要求信号送信部 2030 と、分析部 2040 と、認識部 2050 と、標準パターン記憶部 2060 と、合成音情報取得部 2070 と、圧縮方式インデックス取得部 2080 と、認識エンジン選択部 2090 と、認識エンジン設定部 2100 と、圧縮方式取得部 2110 と、圧縮方式記憶部 2120 とを備えている。

【0142】サーバ側では、端末側の波形送信部 1840 から出力された圧縮波形データ、又は始端キャンセル

信号送信部 1860 から出力された始端キャンセル時の信号、又は圧縮方式要求信号送信部 1920 から出力された波形圧縮方式要求信号を波形・信号受信部 2010 で受信する。波形・信号受信部 2010 では、波形圧縮方式要求信号が受信された場合、受信信号を圧縮方式取得部 2110 へ送る。

【0143】圧縮方式取得部 2110 では、波形・信号受信部 2010 から波形圧縮方式要求信号が送られた場合、圧縮方式記憶部 2120 に記憶されている圧縮方式の中からサーバ側で利用可能な圧縮方式を取得し、取得した圧縮方式情報を端末側に出力する。波形・信号受信部 2010 において波形データが受信された場合、波形伸張部 2020 に送信する。波形データの受信に失敗した場合は、波形データ再送要求信号送信部 2030 に波形データ再送要求信号を送信する。波形データ再送要求信号送信部 2030 では、波形・信号受信部 2010 から送られた信号を出力する。

【0144】波形伸張部 2020 では、波形・信号受信部 2010 から送られたデータを波形データとその他のデータとに分離し、波形データを伸張する。その他の情報は合成音情報取得部 2070 と圧縮方式インデックス取得部 2080 に送られ、伸張された波形データは分析部 2040 に送られる。

【0145】合成音情報取得部 2070 では、波形伸張部 2020 から送られた情報から合成音情報を取得する。取得された合成音情報は認識部 2050 に送られる。

【0146】圧縮方式インデックス取得部 2080 では、波形伸張部 2020 から送られた情報から圧縮方式のインデックスを取得する。

【0147】取得されたインデックスは認識エンジン選択部 2090 に送られる。認識エンジン選択部 2090 では、圧縮方式インデックス取得部 2080 で取得されたインデックスをもとに認識に用いる認識エンジンを選択する。選択されたエンジンの情報は認識エンジン設定部 2100 と標準パターン記憶部 2060 に送られる。

【0148】認識エンジン設定部 2100 では、認識エンジン選択部 2090 から送られたエンジン情報を元に使用するエンジンの分析手法を分析部 2040 へ送り、認識手法を認識部 2050 へ送る。標準パターン記憶部 2060 では、認識エンジン選択部 2090 から送られたエンジン情報より使用する標準パターンを設定する。分析部 2040 では、波形伸張部 2020 で伸張された波形データを特徴ベクトルに変換する。特徴ベクトルとしてはケプストラム、 Δ ケプストラム等が用いられる。分析された特徴ベクトルは認識部 2050 に送られる。認識部 2050 では、分析部 2040 から送られた特徴ベクトルと、標準パターン記憶部 2060 から送られた標準パターンとの間の尤度を計算し、認識結果を求める。尤度の計算方法は Viterbi アルゴリズム、DP マッ

チング等が用いられる。求められた認識結果と合成音情報取得部 2070 で取得された合成音情報は共に端末側に出力されるか、又は、両者が関連付けられることでどの合成音に対する認識結果であるかが対応付けされ、対応付けされた後の認識結果が端末側に出力される。

【0149】波形・信号受信部 2010 に始端キャンセル信号送信部 1860 から送信された始端キャンセル時の信号が入力された場合、信号は認識部 2050 に送られる。認識部 2050 では、波形・信号受信部 2010 から始端キャンセル時の信号が送られてきた場合、認識

を途中で中断する。

【0150】上記実施例において、上記端末側の装置としては、例えば携帯電話端末が好適とされるが、端末は、携帯電話端末に限定されるものではなく、サーバ・クライアント型音声認識装置のサーバに接続可能とされる音声入力可能な任意の情報端末を含む。以上本発明を上記各実施例に即して説明したが、本発明は、上記実施例の構成にのみ限定されるものではなく、各実施例の組合せを含むものであり、本発明は、特許請求の範囲の請求項の範囲内で当業者であればなし得るであろう各種変形、修正を含むことは勿論である。

【0151】

【発明の効果】以上説明したように、本発明によれば、認識対象となるタスクの規模や難易度に適応し、クライアントとサーバ間で相互の音声処理仕様、能力を確認して音声圧縮伸張等の音声処理仕様を最適に設定することができる、という効果を奏する。

【0152】本発明によれば、更に、サーバで認識処理を行うため大規模なタスクの認識処理が可能となり、認識エンジン等のアップデートに迅速に対応することができるという効果を奏する。

【図面の簡単な説明】

【図 1】 (a) は、本発明の第 1 の実施例における端末側の装置構成、(b) は、本発明の第 1 の実施例におけるサーバ側の装置構成を示す図である。

【図 2】 (a) は、本発明の第 2 の実施例における端末側の装置構成、(b) は、本発明の第 2 の実施例におけるサーバ側の装置構成を示す図である。

【図 3】 (a) は、本発明の第 3 の実施例における端末側の装置構成、(b) は、本発明の第 3 の実施例におけるサーバ側の装置構成を示す図である。

【図 4】 (a) は、本発明の第 4 の実施例における端末側の装置構成、(b) は、本発明の第 4 の実施例におけるサーバ側の装置構成を示す図である。

【図 5】 (a) は、本発明の第 5 の実施例における端末側の装置構成、(b) は、本発明の第 5 の実施例におけるサーバ側の装置構成を示す図である。

【図 6】 (a) は、本発明の第 6 の実施例における端末側の装置構成、(b) は、本発明の第 6 の実施例におけるサーバ側の装置構成を示す図である。

【図 7】 (a) は、本発明の第 7 の実施例における端末側の装置構成、(b) は、本発明の第 7 の実施例におけるサーバ側の装置構成を示す図である。

【符号の説明】

10 音声検出部
20 波形圧縮部
30 波形送信部
110 波形受信部
120 波形伸張部
130 分析部
140 認識部
150 標準パターン記憶部
210 波形・信号受信部
220 音声検出部
230 波形圧縮部
240 波形送信部
250 波形記憶部
310 波形受信部
320 波形伸張部
330 波形データ再送要求信号送信部
340 分析部
350 認識部
360 標準パターン記憶部
410 波形・信号受信部
420 音声検出部
430 波形圧縮部
440 波形送信部
450 波形記憶部
460 始端キャンセル信号送信部
510 波形・信号受信部
520 波形伸張部
530 波形データ再送要求信号送信部
540 分析部
550 認識部
560 標準パターン記憶部
610 波形・信号受信部
620 音声検出部
630 波形圧縮部
640 波形送信部
650 波形記憶部
660 始端キャンセル信号送信部
670 音声合成部
680 合成音情報作成部
710 波形・信号受信部
720 波形伸張部
730 波形データ再送要求信号送信部
740 分析部
750 認識部
760 標準パターン記憶部
770 合成音情報取得部

1010 波形・信号・圧縮方式受信部
 1020 音声検出部
 1030 波形圧縮部
 1040 波形送信部
 1050 波形記憶部
 1060 始端キャンセル信号送信部
 1070 音声合成部
 1080 合成音情報作成部
 1090 圧縮方式インデックス作成部
 1100 圧縮方式選択部
 1110 圧縮方式記憶部
 1120 圧縮方式要求信号送信部
 1210 波形・信号受信部
 1220 波形伸張部
 1230 波形データ再送要求信号送信部
 1240 分析部
 1250 認識部
 1260 標準パターン記憶部
 1270 合成音情報取得部
 1280 圧縮方式インデックス取得部
 1290 認識エンジン選択部
 1300 認識エンジン設定部
 1310 圧縮方式取得部
 1320 圧縮方式記憶部
 1410 波形・信号・圧縮方式受信部
 1420 音声検出部
 1430 波形圧縮部
 1440 波形送信部
 1450 波形記憶部
 1460 始端キャンセル信号送信部
 1470 音声合成部
 1480 合成音情報作成部
 1490 圧縮方式インデックス作成部
 1500 圧縮方式選択部
 1510 圧縮方式記憶部
 1520 圧縮方式要求信号送信部
 1610 波形・信号・タスク情報受信部

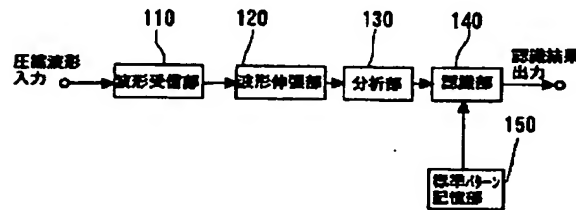
1620 波形伸張部
 1630 波形データ再送要求信号送信部
 1640 分析部
 1650 認識部
 1660 標準パターン記憶部
 1670 合成音情報取得部
 1680 圧縮方式インデックス取得部
 1690 認識エンジン選択部
 1700 認識エンジン設定部
 10 1710 圧縮方式取得部
 1720 圧縮方式・タスク情報対応テーブル記憶部
 1730 タスク情報記憶部
 1810 波形・信号・圧縮方式・タスク情報受信部
 1820 音声検出部
 1830 波形圧縮部
 1840 波形送信部
 1850 波形記憶部
 1860 始端キャンセル信号送信部
 1870 音声合成部
 20 1880 合成音情報作成部
 1890 圧縮方式インデックス作成部
 1900 圧縮方式選択部
 1910 圧縮方式・タスク情報対応テーブル記憶部
 1920 圧縮方式要求信号送信部
 1930 タスク情報記憶部
 2010 波形・信号受信部
 2020 波形伸張部
 2030 波形データ再送要求信号送信部
 2040 分析部
 30 2050 認識部
 2060 標準パターン記憶部
 2070 合成音情報取得部
 2080 圧縮方式インデックス取得部
 2090 認識エンジン選択部
 2100 認識エンジン設定部
 2110 圧縮方式取得部
 2120 圧縮方式記憶部

【図 1】

(a)
【端末側】

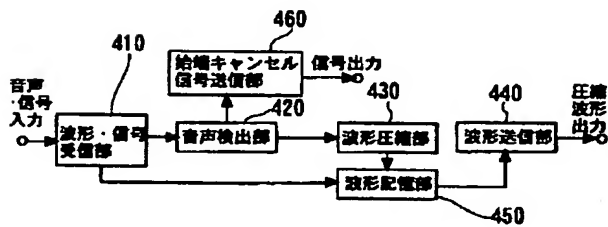


(b)
【サーバ側】

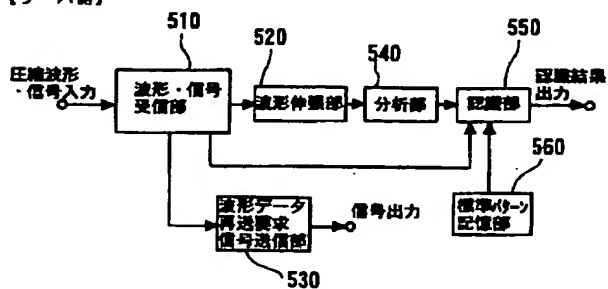


【図 3】

(a)
【端末側】

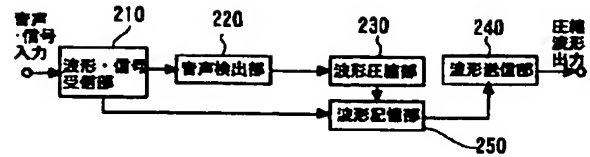


(b)
【サーバ側】

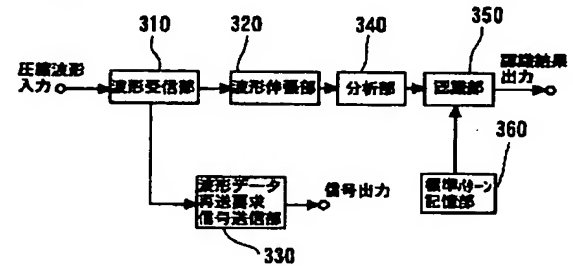


【図 2】

(a)
【端末側】

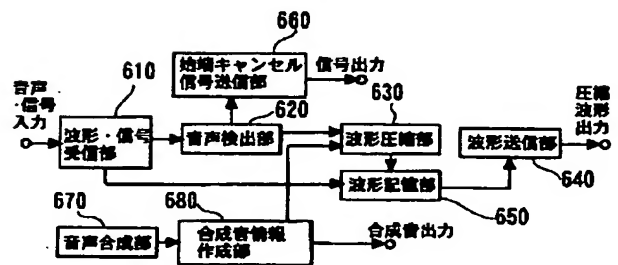


(b)
【サーバ側】

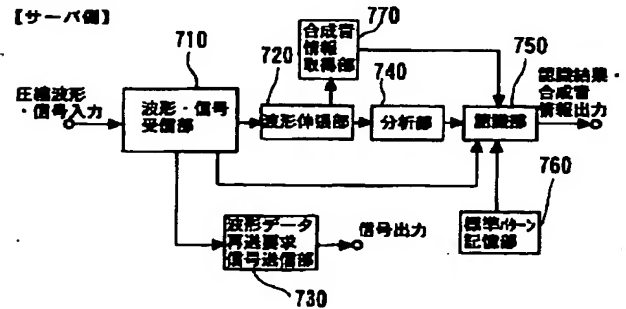


【図 4】

(a)
【端末側】



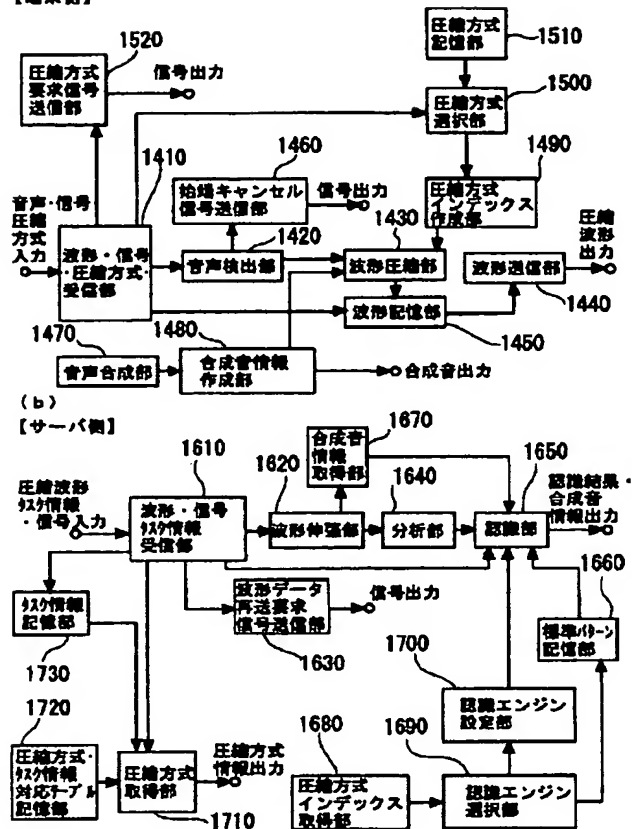
(b)
【サーバ側】



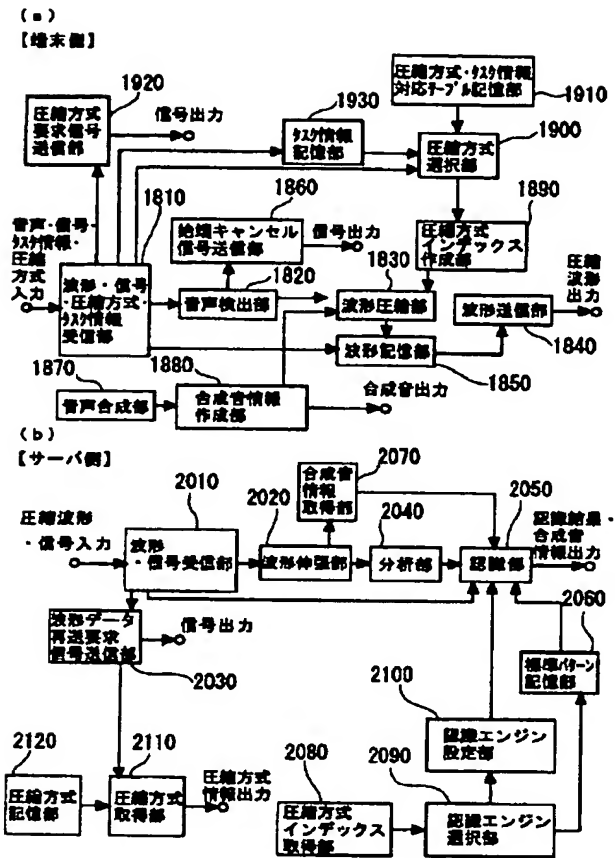
【図 6】

(a)

【端末側】



【図 7】



フロントページの続き

(51) Int. Cl. 7

識別記号

F I

テーマコード* (参考)

G 1 0 L 15/04

(72) 発明者 吉田 和永

東京都港区芝五丁目7番1号 日本電気株式会社内

F ターム (参考) 5D015 DD03 FF01 KK01